Customer_311Project

July 25, 2023

[1]:

```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline
import scipy as sc
from scipy import stats
import warnings
warnings.filterwarnings("ignore")
```

0.1 Understand the dataset

1.1 import the dataset 1.2 visualize the dataset 1.3 print the columns of the dataframe 1.4 Identify the shape of the dataset 1.5 Identify the variable with null values

[2]:

```
service_request =pd.read_csv("311_Service_Requests_from_2010_to_Present.csv")
service_request.head()
```

[2]:

```
   Unique Key            Created Date            Closed Date Agency  \
0    32310363  12/31/2015 11:59:45 PM  01/01/2016 12:55:15 AM   NYPD
1    32309934  12/31/2015 11:59:44 PM  01/01/2016 01:26:57 AM   NYPD
2    32309159  12/31/2015 11:59:29 PM  01/01/2016 04:51:03 AM   NYPD
3    32305098  12/31/2015 11:57:46 PM  01/01/2016 07:43:13 AM   NYPD
4    32306529  12/31/2015 11:56:58 PM  01/01/2016 03:24:42 AM   NYPD

                       Agency Name           Complaint Type  \
0  New York City Police Department  Noise - Street/Sidewalk
1  New York City Police Department         Blocked Driveway
2  New York City Police Department         Blocked Driveway
3  New York City Police Department          Illegal Parking
4  New York City Police Department          Illegal Parking

                    Descriptor    Location Type  Incident Zip  \
0             Loud Music/Party  Street/Sidewalk       10034.0
1                    No Access  Street/Sidewalk       11105.0
2                    No Access  Street/Sidewalk       10458.0
3  Commercial Overnight Parking  Street/Sidewalk       10461.0
```

```
4              Blocked Sidewalk  Street/Sidewalk       11373.0

        Incident Address  … Bridge Highway Name Bridge Highway Direction  \
0    71 VERMILYEA AVENUE  …                 NaN                      NaN
1        27-07 23 AVENUE  …                 NaN                      NaN
2  2897 VALENTINE AVENUE  …                 NaN                      NaN
3    2940 BAISLEY AVENUE  …                 NaN                      NaN
4          87-14 57 ROAD  …                 NaN                      NaN

  Road Ramp Bridge Highway Segment Garage Lot Name Ferry Direction  \
0       NaN                    NaN             NaN             NaN
1       NaN                    NaN             NaN             NaN
2       NaN                    NaN             NaN             NaN
3       NaN                    NaN             NaN             NaN
4       NaN                    NaN             NaN             NaN

  Ferry Terminal Name   Latitude  Longitude  \
0                 NaN  40.865682 -73.923501
1                 NaN  40.775945 -73.915094
2                 NaN  40.870325 -73.888525
3                 NaN  40.835994 -73.828379
4                 NaN  40.733060 -73.874170

                                  Location
0   (40.86568153633767, -73.92350095571744)
1  (40.775945312321085, -73.91509393898605)
2  (40.870324522111424, -73.88852464418646)
3   (40.83599404683083, -73.82837939584206)
4  (40.733059618956815, -73.87416975810375)

[5 rows x 53 columns]
```

[3]:
```
service_request.tail()
```

[3]:
```
        Unique Key            Created Date             Closed Date Agency  \
364553    29609918  01/01/2015 12:04:44 AM  01/01/2015 10:22:31 AM   NYPD
364554    29608392  01/01/2015 12:04:28 AM  01/01/2015 02:25:02 AM   NYPD
364555    29607589  01/01/2015 12:01:30 AM  01/01/2015 12:20:33 AM   NYPD
364556    29610889  01/01/2015 12:01:29 AM  01/01/2015 02:42:22 AM   NYPD
364557    29611816  01/01/2015 12:00:50 AM  01/01/2015 02:47:50 AM   NYPD

                            Agency Name           Complaint Type  \
364553  New York City Police Department          Illegal Parking
364554  New York City Police Department          Noise - Vehicle
364555  New York City Police Department  Noise - Street/Sidewalk
364556  New York City Police Department         Blocked Driveway
364557  New York City Police Department         Blocked Driveway
```

```
              Descriptor    Location Type  Incident Zip      Incident Address  \
364553   Blocked Hydrant  Street/Sidewalk       11421.0         84-25 85 ROAD
364554    Car/Truck Horn  Street/Sidewalk       10468.0  2555 SEDGWICK AVENUE
364555  Loud Music/Party  Street/Sidewalk       10031.0   508 WEST 139 STREET
364556         No Access  Street/Sidewalk       10466.0   931 EAST 226 STREET
364557         No Access  Street/Sidewalk       11420.0     123-19 135 STREET

        … Bridge Highway Name Bridge Highway Direction Road Ramp  \
364553  …                 NaN                      NaN       NaN
364554  …                 NaN                      NaN       NaN
364555  …                 NaN                      NaN       NaN
364556  …                 NaN                      NaN       NaN
364557  …                 NaN                      NaN       NaN

       Bridge Highway Segment Garage Lot Name Ferry Direction  \
364553                    NaN             NaN             NaN
364554                    NaN             NaN             NaN
364555                    NaN             NaN             NaN
364556                    NaN             NaN             NaN
364557                    NaN             NaN             NaN

       Ferry Terminal Name   Latitude  Longitude  \
364553                 NaN  40.695145 -73.860949
364554                 NaN  40.867830 -73.907178
364555                 NaN  40.821647 -73.950873
364556                 NaN  40.886361 -73.853290
364557                 NaN  40.674212 -73.803585

                                        Location
364553   (40.69514470265117, -73.86094888534394)
364554   (40.86782963689454, -73.90717786644662)
364555  (40.821646626438095, -73.95087342885292)
364556   (40.88636077906953, -73.85329048666742)
364557  (40.674211762243935, -73.80358548685278)

[5 rows x 53 columns]
```

[4]:

```
#to check the columns
service_request.columns
```

[4]:

```
Index(['Unique Key', 'Created Date', 'Closed Date', 'Agency', 'Agency Name',
       'Complaint Type', 'Descriptor', 'Location Type', 'Incident Zip',
       'Incident Address', 'Street Name', 'Cross Street 1', 'Cross Street 2',
       'Intersection Street 1', 'Intersection Street 2', 'Address Type',
       'City', 'Landmark', 'Facility Type', 'Status', 'Due Date',
       'Resolution Description', 'Resolution Action Updated Date',
       'Community Board', 'Borough', 'X Coordinate (State Plane)',
       'Y Coordinate (State Plane)', 'Park Facility Name', 'Park Borough',
       'School Name', 'School Number', 'School Region', 'School Code',
       'School Phone Number', 'School Address', 'School City', 'School State',
       'School Zip', 'School Not Found', 'School or Citywide Complaint',
       'Vehicle Type', 'Taxi Company Borough', 'Taxi Pick Up Location',
       'Bridge Highway Name', 'Bridge Highway Direction', 'Road Ramp',
       'Bridge Highway Segment', 'Garage Lot Name', 'Ferry Direction',
       'Ferry Terminal Name', 'Latitude', 'Longitude', 'Location'],
      dtype='object')
```

```
[5]: service_request.shape
```

```
[5]: (364558, 53)
```

```
[6]: service_request.describe
```

```
[6]: <bound method NDFrame.describe of         Unique Key            Created Date
     Closed Date Agency  \
     0         32310363  12/31/2015 11:59:45 PM  01/01/2016 12:55:15 AM   NYPD
     1         32309934  12/31/2015 11:59:44 PM  01/01/2016 01:26:57 AM   NYPD
     2         32309159  12/31/2015 11:59:29 PM  01/01/2016 04:51:03 AM   NYPD
     3         32305098  12/31/2015 11:57:46 PM  01/01/2016 07:43:13 AM   NYPD
     4         32306529  12/31/2015 11:56:58 PM  01/01/2016 03:24:42 AM   NYPD
     …              …                      …                      …    …
     364553    29609918  01/01/2015 12:04:44 AM  01/01/2015 10:22:31 AM   NYPD
     364554    29608392  01/01/2015 12:04:28 AM  01/01/2015 02:25:02 AM   NYPD
     364555    29607589  01/01/2015 12:01:30 AM  01/01/2015 12:20:33 AM   NYPD
     364556    29610889  01/01/2015 12:01:29 AM  01/01/2015 02:42:22 AM   NYPD
     364557    29611816  01/01/2015 12:00:50 AM  01/01/2015 02:47:50 AM   NYPD

                                 Agency Name           Complaint Type  \
     0       New York City Police Department  Noise - Street/Sidewalk
     1       New York City Police Department         Blocked Driveway
     2       New York City Police Department         Blocked Driveway
     3       New York City Police Department          Illegal Parking
     4       New York City Police Department          Illegal Parking
     …                                   …                      …
     364553  New York City Police Department          Illegal Parking
     364554  New York City Police Department          Noise - Vehicle
     364555  New York City Police Department  Noise - Street/Sidewalk
     364556  New York City Police Department         Blocked Driveway
     364557  New York City Police Department         Blocked Driveway

                               Descriptor    Location Type  Incident Zip  \
     0                   Loud Music/Party  Street/Sidewalk       10034.0
     1                          No Access  Street/Sidewalk       11105.0
```

```
2                            No Access  Street/Sidewalk       10458.0
3       Commercial Overnight Parking  Street/Sidewalk       10461.0
4                   Blocked Sidewalk  Street/Sidewalk       11373.0
…                               …              …           …
364553               Blocked Hydrant  Street/Sidewalk       11421.0
364554                Car/Truck Horn  Street/Sidewalk       10468.0
364555              Loud Music/Party  Street/Sidewalk       10031.0
364556                     No Access  Street/Sidewalk       10466.0
364557                     No Access  Street/Sidewalk       11420.0

             Incident Address  … Bridge Highway Name  \
0         71 VERMILYEA AVENUE  …                NaN
1             27-07 23 AVENUE  …                NaN
2       2897 VALENTINE AVENUE  …                NaN
3         2940 BAISLEY AVENUE  …                NaN
4               87-14 57 ROAD  …                NaN
…                        …  …                 …
364553          84-25 85 ROAD  …                NaN
364554   2555 SEDGWICK AVENUE  …                NaN
364555    508 WEST 139 STREET  …                NaN
364556    931 EAST 226 STREET  …                NaN
364557      123-19 135 STREET  …                NaN

       Bridge Highway Direction Road Ramp Bridge Highway Segment  \
0                           NaN       NaN                    NaN
1                           NaN       NaN                    NaN
2                           NaN       NaN                    NaN
3                           NaN       NaN                    NaN
4                           NaN       NaN                    NaN
…                         …       …                     …
364553                      NaN       NaN                    NaN
364554                      NaN       NaN                    NaN
364555                      NaN       NaN                    NaN
364556                      NaN       NaN                    NaN
364557                      NaN       NaN                    NaN

       Garage Lot Name Ferry Direction Ferry Terminal Name   Latitude  \
0                  NaN             NaN                 NaN  40.865682
1                  NaN             NaN                 NaN  40.775945
2                  NaN             NaN                 NaN  40.870325
3                  NaN             NaN                 NaN  40.835994
4                  NaN             NaN                 NaN  40.733060
…                 …             …                 …        …
364553             NaN             NaN                 NaN  40.695145
364554             NaN             NaN                 NaN  40.867830
364555             NaN             NaN                 NaN  40.821647
364556             NaN             NaN                 NaN  40.886361
```

```
364557             NaN             NaN                 NaN  40.674212

        Longitude                                   Location
0      -73.923501   (40.86568153633767, -73.92350095571744)
1      -73.915094  (40.775945312321085, -73.91509393898605)
2      -73.888525  (40.870324522111424, -73.88852464418646)
3      -73.828379   (40.83599404683083, -73.82837939584206)
4      -73.874170  (40.733059618956815, -73.87416975810375)
…           …                                           …
364553 -73.860949   (40.69514470265117, -73.86094888534394)
364554 -73.907178   (40.86782963689454, -73.90717786644662)
364555 -73.950873  (40.821646626438095, -73.95087342885292)
364556 -73.853290   (40.88636077906953, -73.85329048666742)
364557 -73.803585  (40.674211762243935, -73.80358548685278)

[364558 rows x 53 columns]>
```

[7]:
```
#to check the information
service_request.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 364558 entries, 0 to 364557
Data columns (total 53 columns):
 #   Column                          Non-Null Count   Dtype
---  ------                          --------------   -----
 0   Unique Key                      364558 non-null  int64
 1   Created Date                    364558 non-null  object
 2   Closed Date                     362177 non-null  object
 3   Agency                          364558 non-null  object
 4   Agency Name                     364558 non-null  object
 5   Complaint Type                  364558 non-null  object
 6   Descriptor                      358057 non-null  object
 7   Location Type                   364425 non-null  object
 8   Incident Zip                    361560 non-null  float64
 9   Incident Address                312859 non-null  object
 10  Street Name                     312859 non-null  object
 11  Cross Street 1                  307370 non-null  object
 12  Cross Street 2                  306753 non-null  object
 13  Intersection Street 1           51120 non-null   object
 14  Intersection Street 2           50512 non-null   object
 15  Address Type                    361306 non-null  object
 16  City                            361561 non-null  object
 17  Landmark                        375 non-null     object
 18  Facility Type                   362169 non-null  object
 19  Status                          364558 non-null  object
 20  Due Date                        364555 non-null  object
 21  Resolution Description          364558 non-null  object
 22  Resolution Action Updated Date  362156 non-null  object
 23  Community Board                 364558 non-null  object
 24  Borough                         364558 non-null  object
 25  X Coordinate (State Plane)      360528 non-null  float64
 26  Y Coordinate (State Plane)      360528 non-null  float64
 27  Park Facility Name              364558 non-null  object
 28  Park Borough                    364558 non-null  object
 29  School Name                     364558 non-null  object
 30  School Number                   364558 non-null  object
 31  School Region                   364557 non-null  object
 32  School Code                     364557 non-null  object
 33  School Phone Number             364558 non-null  object
 34  School Address                  364558 non-null  object
 35  School City                     364558 non-null  object
 36  School State                    364558 non-null  object
 37  School Zip                      364557 non-null  object
 38  School Not Found                364558 non-null  object
 39  School or Citywide Complaint    0 non-null       float64
 40  Vehicle Type                    0 non-null       float64
 41  Taxi Company Borough            0 non-null       float64
 42  Taxi Pick Up Location           0 non-null       float64
 43  Bridge Highway Name             297 non-null     object
 44  Bridge Highway Direction        297 non-null     object
 45  Road Ramp                       262 non-null     object
 46  Bridge Highway Segment          262 non-null     object
 47  Garage Lot Name                 0 non-null       float64
 48  Ferry Direction                 1 non-null       object
 49  Ferry Terminal Name             2 non-null       object
 50  Latitude                        360528 non-null  float64
 51  Longitude                       360528 non-null  float64
 52  Location                        360528 non-null  object
dtypes: float64(10), int64(1), object(42)
memory usage: 147.4+ MB
```

[8]:
```
# to check the null counts in a rows
service_request.isna().sum(axis=0)
```

[8]:
```
Unique Key                            0
Created Date                          0
Closed Date                        2381
Agency                                0
Agency Name                           0
Complaint Type                        0
Descriptor                         6501
Location Type                       133
Incident Zip                       2998
Incident Address                  51699
```

```
Street Name                        51699
Cross Street 1                     57188
Cross Street 2                     57805
Intersection Street 1             313438
Intersection Street 2             314046
Address Type                        3252
City                                2997
Landmark                          364183
Facility Type                       2389
Status                                 0
Due Date                               3
Resolution Description                 0
Resolution Action Updated Date      2402
Community Board                        0
Borough                                0
X Coordinate (State Plane)          4030
Y Coordinate (State Plane)          4030
Park Facility Name                     0
Park Borough                           0
School Name                            0
School Number                          0
School Region                          1
School Code                            1
School Phone Number                    0
School Address                         0
School City                            0
School State                           0
School Zip                             1
School Not Found                       0
School or Citywide Complaint      364558
Vehicle Type                      364558
Taxi Company Borough              364558
Taxi Pick Up Location             364558
Bridge Highway Name               364261
Bridge Highway Direction          364261
Road Ramp                         364296
Bridge Highway Segment            364296
Garage Lot Name                   364558
Ferry Direction                   364557
Ferry Terminal Name               364556
Latitude                            4030
Longitude                           4030
Location                            4030
dtype: int64
```

[9]:
```
service_request.isna().sum(axis=1).value_counts()
```

```
[9]: 14    300934
     16     51856
     15      4935
     17      2252
     27      1832
     19      1014
     18       515
     24       500
     25       482
     22       122
     20        91
     28        15
     21         3
     26         3
     23         2
     29         1
     13         1
     dtype: int64
```

0.2 Perform Basic DATA EXPLORATORY ANALYSIS:

2.1 draw frequency plot to show the number of null values in each column of the DataFrames. 2.2 Missing Value Treatment 2.2.1 Remove the record whose ”closed Date” values are null

```
[10]: Service_data_null = service_request.isna().sum(axis=0)
      Service_data_null
```

```
[10]: Unique Key                             0
      Created Date                           0
      Closed Date                         2381
      Agency                                 0
      Agency Name                            0
      Complaint Type                         0
      Descriptor                          6501
      Location Type                        133
      Incident Zip                        2998
      Incident Address                   51699
      Street Name                        51699
      Cross Street 1                     57188
      Cross Street 2                     57805
      Intersection Street 1             313438
      Intersection Street 2             314046
      Address Type                        3252
      City                                2997
      Landmark                          364183
      Facility Type                       2389
```

```
Status                                 0
Due Date                               3
Resolution Description                 0
Resolution Action Updated Date      2402
Community Board                        0
Borough                                0
X Coordinate (State Plane)          4030
Y Coordinate (State Plane)          4030
Park Facility Name                     0
Park Borough                           0
School Name                            0
School Number                          0
School Region                          1
School Code                            1
School Phone Number                    0
School Address                         0
School City                            0
School State                           0
School Zip                             1
School Not Found                       0
School or Citywide Complaint      364558
Vehicle Type                      364558
Taxi Company Borough              364558
Taxi Pick Up Location             364558
Bridge Highway Name               364261
Bridge Highway Direction          364261
Road Ramp                         364296
Bridge Highway Segment            364296
Garage Lot Name                   364558
Ferry Direction                   364557
Ferry Terminal Name               364556
Latitude                            4030
Longitude                           4030
Location                            4030
dtype: int64
```

[11]:

```
service_request.dtypes
```

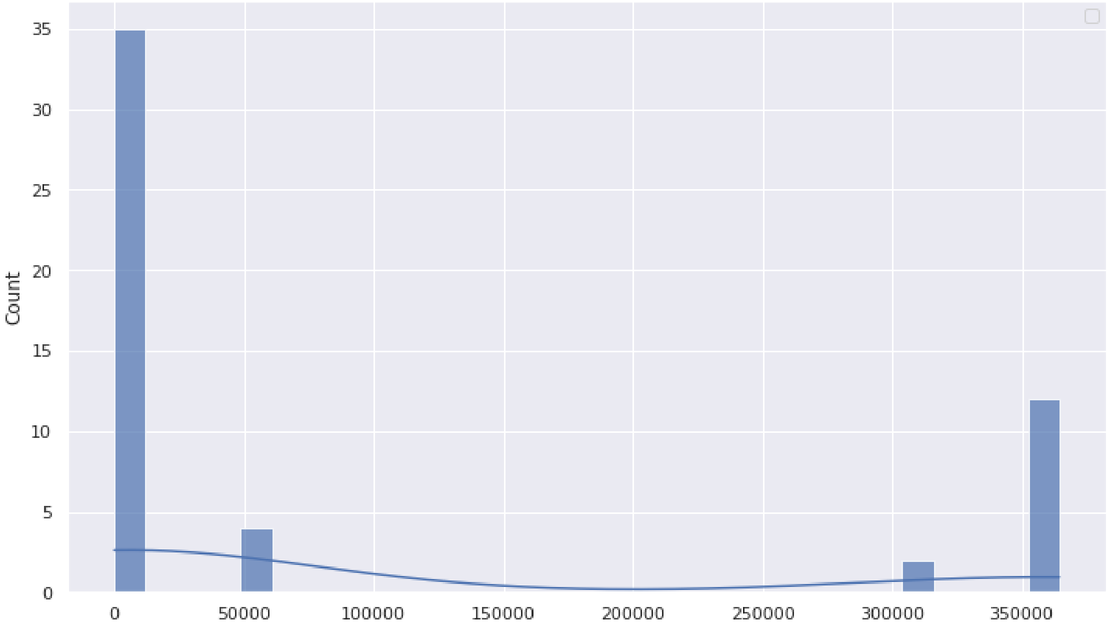
[11]:

```
Unique Key                          int64
Created Date                       object
Closed Date                        object
Agency                             object
Agency Name                        object
Complaint Type                     object
Descriptor                         object
Location Type                      object
Incident Zip                      float64
Incident Address                   object
Street Name                        object
Cross Street 1                     object
Cross Street 2                     object
Intersection Street 1              object
Intersection Street 2              object
Address Type                       object
City                               object
Landmark                           object
Facility Type                      object
Status                             object
Due Date                           object
Resolution Description             object
Resolution Action Updated Date     object
Community Board                    object
Borough                            object
X Coordinate (State Plane)        float64
Y Coordinate (State Plane)        float64
Park Facility Name                 object
Park Borough                       object
School Name                        object
School Number                      object
School Region                      object
School Code                        object
School Phone Number                object
School Address                     object
School City                        object
School State                       object
School Zip                         object
School Not Found                   object
School or Citywide Complaint      float64
Vehicle Type                      float64
Taxi Company Borough              float64
Taxi Pick Up Location             float64
Bridge Highway Name                object
Bridge Highway Direction           object
Road Ramp                          object
Bridge Highway Segment             object
Garage Lot Name                   float64
Ferry Direction                    object
Ferry Terminal Name                object
Latitude                          float64
Longitude                         float64
Location                           object
dtype: object
```

```
[12]: # Draw the frequency plot to show the null values in each dataframe
sns.set(rc={'figure.figsize':(12,7)})
sns.histplot(Service_data_null,
             bins=30,
             kde=True,
             alpha=0.7,
            linewidth =0.8
           )


plt.legend()
```

```
No artists with labels found to put in legend.  Note that artists whose label
start with an underscore are ignored when legend() is called with no argument.
```

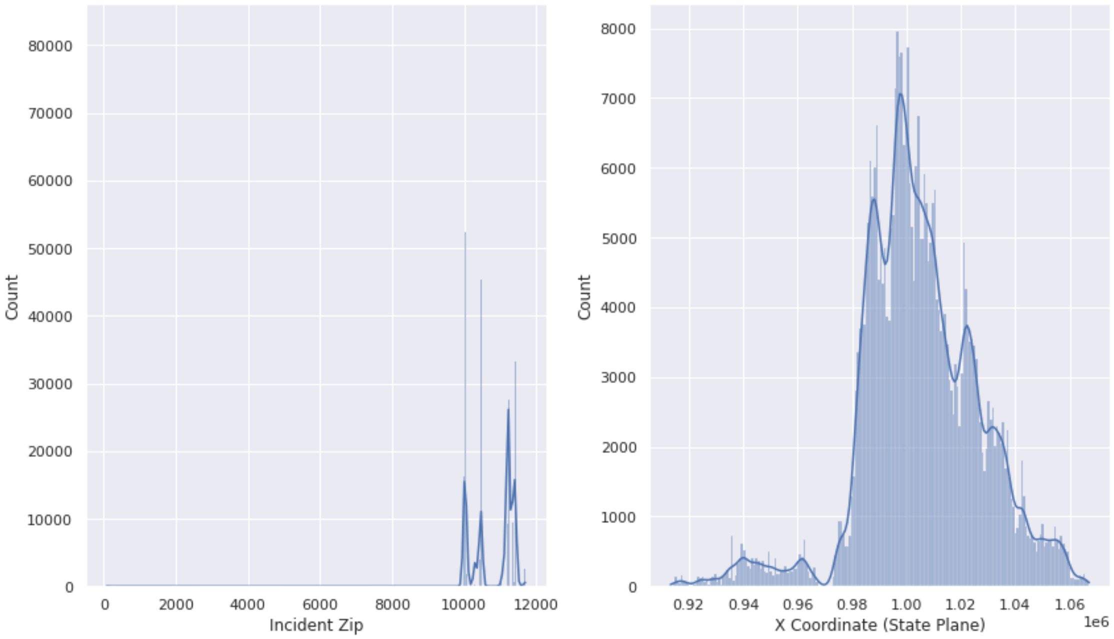
```
[12]: <matplotlib.legend.Legend at 0x7f23776e69d0>
```

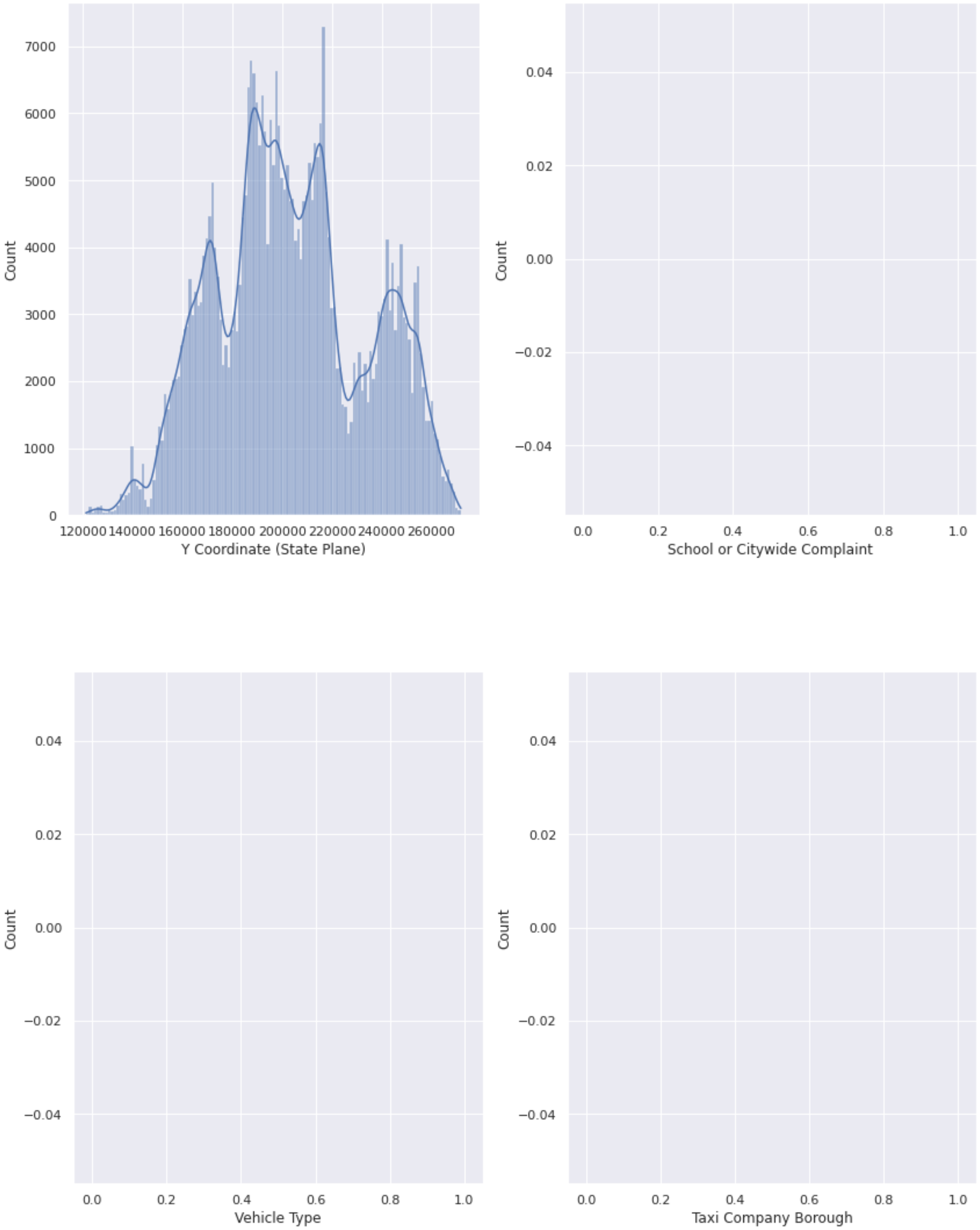


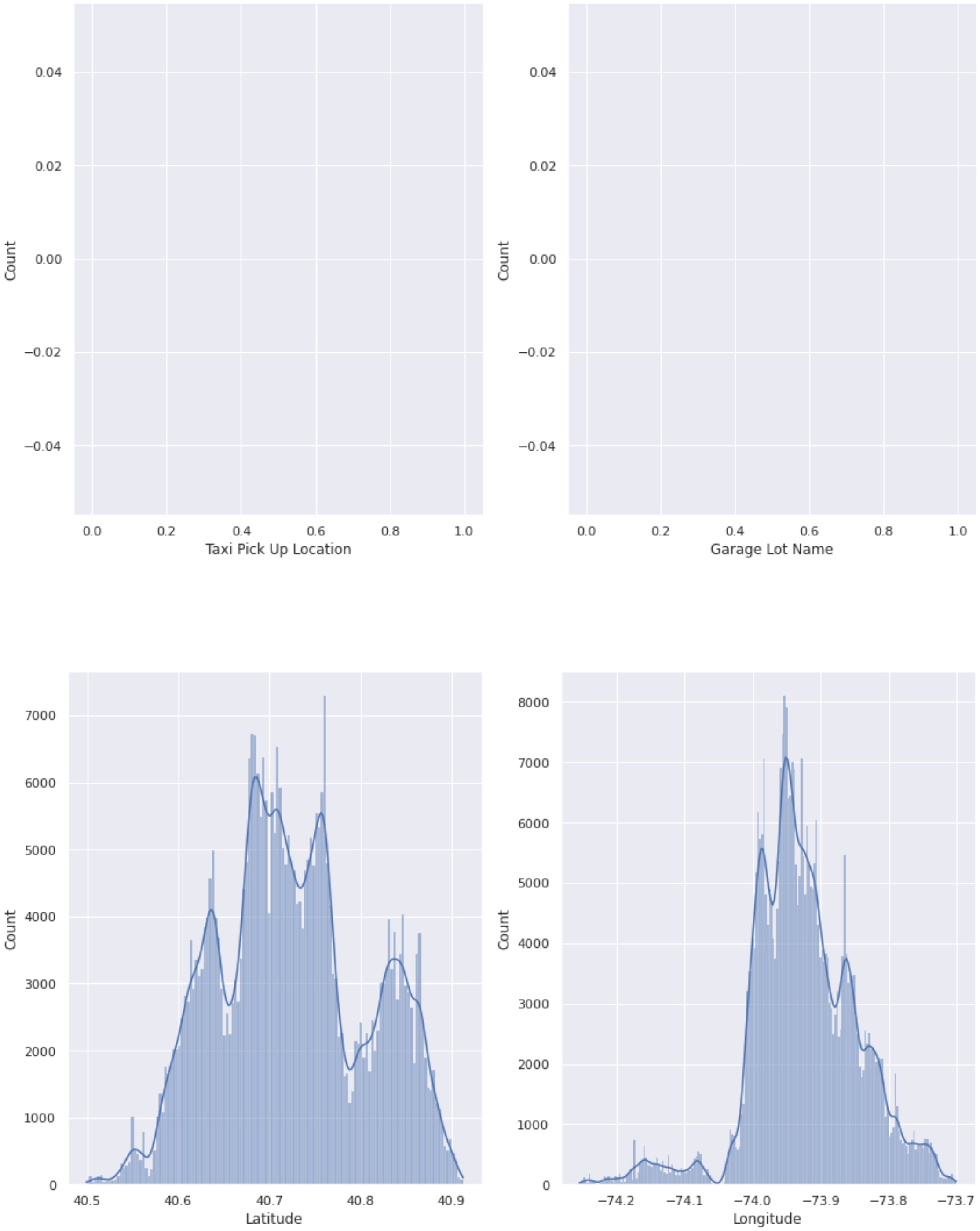
More Clearely for the particular column which are null

```
[13]: import seaborn as sns

num_cols = ['Incident Zip','X Coordinate (State Plane)','Y Coordinate (State␣
 ↪Plane)','School or Citywide Complaint',
            'Vehicle Type','Taxi Company Borough',
            'Taxi Pick Up Location','Garage Lot Name','Latitude','Longitude']

for i in range(0, len(num_cols), 2):
```

```
plt.subplot(121)
sns.histplot(service_request[num_cols[i]],
             multiple="stack",
             kde=True)
plt.subplot(122)
sns.histplot(service_request[num_cols[i+1]],
             multiple="stack",
             kde=True)
plt.tight_layout()
plt.show()
```




Count
7000
6000
5000
4000
3000
2000
1000
0
120000
140000
160000
180000
200000
220000
240000
260000
Y Coordinate (State Plane)
Count
0.04
0.02
0.00
−0.02
−0.04
0.0
0.2
0.4
0.6
0.8
1.0
School or Citywide Complaint
Count
0.04
0.02
0.00
−0.02
−0.04
0.0
0.2
0.4
0.6
0.8
1.0
Vehicle Type
Count
0.04
0.02
0.00
−0.02
−0.04
0.0
0.2
0.4
0.6
0.8
1.0
Taxi Company Borough



we can see in the above graph thatcolumns -’Vehicle Type’,’Taxi Company Borough’, ’Taxi Pick Up Location’,’Garage Lot Name’ are 100% null.

0.3 Missing value treatment

a) Remove the record whose ”closed date” values are null

lets find the missing value percentage so that we can drop more columns it is unnecessary.

[14]:
```
percentage_null_value= (service_request[service_request.columns].isna().
 ↪sum(axis=0)/service_request.shape[0])*100
percentage_null_value
```

[14]:
```
Unique Key                          0.000000
Created Date                        0.000000
Closed Date                         0.653120
Agency                              0.000000
Agency Name                         0.000000
Complaint Type                      0.000000
Descriptor                          1.783255
Location Type                       0.036483
Incident Zip                        0.822366
Incident Address                   14.181283
Street Name                        14.181283
Cross Street 1                     15.686941
Cross Street 2                     15.856187
Intersection Street 1              85.977540
Intersection Street 2              86.144317
Address Type                        0.892039
City                                0.822091
Landmark                           99.897136
Facility Type                       0.655314
Status                              0.000000
Due Date                            0.000823
Resolution Description              0.000000
Resolution Action Updated Date      0.658880
Community Board                     0.000000
Borough                             0.000000
X Coordinate (State Plane)          1.105448
Y Coordinate (State Plane)          1.105448
Park Facility Name                  0.000000
Park Borough                        0.000000
School Name                         0.000000
School Number                       0.000000
School Region                       0.000274
School Code                         0.000274
School Phone Number                 0.000000
School Address                      0.000000
School City                         0.000000
School State                        0.000000
School Zip                          0.000274
School Not Found                    0.000000
School or Citywide Complaint      100.000000
Vehicle Type                      100.000000
Taxi Company Borough              100.000000
Taxi Pick Up Location             100.000000
Bridge Highway Name                99.918531
Bridge Highway Direction           99.918531
Road Ramp                          99.928132
Bridge Highway Segment             99.928132
Garage Lot Name                   100.000000
Ferry Direction                    99.999726
Ferry Terminal Name                99.999451
Latitude                            1.105448
Longitude                           1.105448
Location                            1.105448
dtype: float64
```

[15]:

```
service_request.shape
```

[15]:

```
(364558, 53)
```

[16]:

```
 service_request =service_request.dropna(subset=["Closed Date"],axis=0)
service_request.shape
```

[16]:

```
(362177, 53)
```

[17]:

```
service_request.isna().sum(axis=0)
```

[17]:

```
Unique Key                             0
Created Date                           0
Closed Date                            0
Agency                                 0
Agency Name                            0
Complaint Type                         0
Descriptor                          6496
Location Type                        130
Incident Zip                         675
Incident Address                   51686
Street Name                        51686
Cross Street 1                     55331
Cross Street 2                     55464
Intersection Street 1             311549
Intersection Street 2             311673
Address Type                         929
City                                 674
Landmark                          361802
Facility Type                         18
```

```
Status                                 0
Due Date                               1
Resolution Description                 0
Resolution Action Updated Date        39
Community Board                        0
Borough                                0
X Coordinate (State Plane)          1707
Y Coordinate (State Plane)          1707
Park Facility Name                     0
Park Borough                           0
School Name                            0
School Number                          0
School Region                          1
School Code                            1
School Phone Number                    0
School Address                         0
School City                            0
School State                           0
School Zip                             1
School Not Found                       0
School or Citywide Complaint      362177
Vehicle Type                      362177
Taxi Company Borough              362177
Taxi Pick Up Location             362177
Bridge Highway Name               361880
Bridge Highway Direction          361880
Road Ramp                         361915
Bridge Highway Segment            361915
Garage Lot Name                   362177
Ferry Direction                   362177
Ferry Terminal Name               362177
Latitude                            1707
Longitude                           1707
Location                            1707
dtype: int64
```

0.4 Analyse the date column and remove entries that have incorrect timelines

[18]:
```
print("Created Date:",service_request['Created Date'])
print("Due Date:",service_request['Due Date'])
print("Resolution Action Updated Date:",service_request['Resolution Action␣
 ↪Updated Date'])
print("Closed Date:",service_request['Closed Date'])
```

```
Created Date: 0         12/31/2015 11:59:45 PM
1         12/31/2015 11:59:44 PM
```

```
2         12/31/2015 11:59:29 PM
3         12/31/2015 11:57:46 PM
4         12/31/2015 11:56:58 PM
                   …
364553    01/01/2015 12:04:44 AM
364554    01/01/2015 12:04:28 AM
364555    01/01/2015 12:01:30 AM
364556    01/01/2015 12:01:29 AM
364557    01/01/2015 12:00:50 AM
Name: Created Date, Length: 362177, dtype: object
Due Date: 0         01/01/2016 07:59:45 AM
1         01/01/2016 07:59:44 AM
2         01/01/2016 07:59:29 AM
3         01/01/2016 07:57:46 AM
4         01/01/2016 07:56:58 AM
                   …
364553    01/01/2015 08:04:44 AM
364554    01/01/2015 08:04:28 AM
364555    01/01/2015 08:01:30 AM
364556    01/01/2015 08:01:29 AM
364557    01/01/2015 08:00:50 AM
Name: Due Date, Length: 362177, dtype: object
Resolution Action Updated Date: 0         01/01/2016 12:55:15 AM
1         01/01/2016 01:26:57 AM
2         01/01/2016 04:51:03 AM
3         01/01/2016 07:43:13 AM
4         01/01/2016 03:24:42 AM
                   …
364553    01/01/2015 10:22:31 AM
364554    01/01/2015 02:25:02 AM
364555    01/01/2015 12:20:33 AM
364556    01/01/2015 02:42:22 AM
364557    01/01/2015 02:47:50 AM
Name: Resolution Action Updated Date, Length: 362177, dtype: object
Closed Date: 0         01/01/2016 12:55:15 AM
1         01/01/2016 01:26:57 AM
2         01/01/2016 04:51:03 AM
3         01/01/2016 07:43:13 AM
4         01/01/2016 03:24:42 AM
                   …
364553    01/01/2015 10:22:31 AM
364554    01/01/2015 02:25:02 AM
364555    01/01/2015 12:20:33 AM
364556    01/01/2015 02:42:22 AM
364557    01/01/2015 02:47:50 AM
Name: Closed Date, Length: 362177, dtype: object
```

```
[19]: service_request['Created Date'] = pd.to_datetime(service_request['Created␣
       ↪Date'],errors='coerce')
      print("Created Date:",service_request['Created Date'])
```

```
Created Date: 0         2015-12-31 23:59:45
1         2015-12-31 23:59:44
2         2015-12-31 23:59:29
3         2015-12-31 23:57:46
4         2015-12-31 23:56:58
                  …
364553    2015-01-01 00:04:44
364554    2015-01-01 00:04:28
364555    2015-01-01 00:01:30
364556    2015-01-01 00:01:29
364557    2015-01-01 00:00:50
Name: Created Date, Length: 362177, dtype: datetime64[ns]
```

```
[20]: # Conversion into datetime
      service_request['Due Date'] = pd.to_datetime(service_request['Due␣
       ↪Date'],errors='coerce')
      print("Due Date:",service_request['Due Date'])
```

```
Due Date: 0         2016-01-01 07:59:45
1         2016-01-01 07:59:44
2         2016-01-01 07:59:29
3         2016-01-01 07:57:46
4         2016-01-01 07:56:58
                  …
364553    2015-01-01 08:04:44
364554    2015-01-01 08:04:28
364555    2015-01-01 08:01:30
364556    2015-01-01 08:01:29
364557    2015-01-01 08:00:50
Name: Due Date, Length: 362177, dtype: datetime64[ns]
```

```
[21]: service_request['Resolution Action Updated Date'] = pd.
       ↪to_datetime(service_request['Resolution Action Updated␣
       ↪Date'],errors='coerce')
      print("Resolution Action Updated Date:",service_request['Resolution Action␣
       ↪Updated Date'])
```

```
Resolution Action Updated Date: 0         2016-01-01 00:55:15
1         2016-01-01 01:26:57
2         2016-01-01 04:51:03
3         2016-01-01 07:43:13
4         2016-01-01 03:24:42
                  …
364553    2015-01-01 10:22:31
```

```
364554   2015-01-01 02:25:02
364555   2015-01-01 00:20:33
364556   2015-01-01 02:42:22
364557   2015-01-01 02:47:50
Name: Resolution Action Updated Date, Length: 362177, dtype: datetime64[ns]
```

```
[ ]: service_request['Closed Date'] = pd.to_datetime(service_request['Closed␣
      ↪Date'],errors='coerce')
     print("Closed Date:",service_request['Closed Date'])
```

In the above we hve seen that the dtype object is converted into datetime64[ns], and the invalid entries is being replaced by NAT(not a time) value to easy access.

```
[73]: service_request.tail()
```

```
[73]:         Unique Key        Created Date         Closed Date Agency  \
       364553   29609918 2015-01-01 00:04:44 2015-01-01 10:22:31   NYPD
       364554   29608392 2015-01-01 00:04:28 2015-01-01 02:25:02   NYPD
       364555   29607589 2015-01-01 00:01:30 2015-01-01 00:20:33   NYPD
       364556   29610889 2015-01-01 00:01:29 2015-01-01 02:42:22   NYPD
       364557   29611816 2015-01-01 00:00:50 2015-01-01 02:47:50   NYPD

                                  Agency Name           Complaint Type  \
       364553  New York City Police Department          Illegal Parking
       364554  New York City Police Department          Noise - Vehicle
       364555  New York City Police Department  Noise - Street/Sidewalk
       364556  New York City Police Department         Blocked Driveway
       364557  New York City Police Department         Blocked Driveway

                     Descriptor    Location Type  Incident Zip      Incident Address  \
       364553   Blocked Hydrant  Street/Sidewalk       11421.0         84-25 85 ROAD
       364554    Car/Truck Horn  Street/Sidewalk       10468.0  2555 SEDGWICK AVENUE
       364555  Loud Music/Party  Street/Sidewalk       10031.0    508 WEST 139 STREET
       364556         No Access  Street/Sidewalk       10466.0    931 EAST 226 STREET
       364557         No Access  Street/Sidewalk       11420.0     123-19 135 STREET

               … Road Ramp Bridge Highway Segment Garage Lot Name Ferry Direction  \
       364553  …       NaN                    NaN             NaN             NaN
       364554  …       NaN                    NaN             NaN             NaN
       364555  …       NaN                    NaN             NaN             NaN
       364556  …       NaN                    NaN             NaN             NaN
       364557  …       NaN                    NaN             NaN             NaN

              Ferry Terminal Name   Latitude  Longitude  \
       364553                 NaN  40.695145 -73.860949
       364554                 NaN  40.867830 -73.907178
       364555                 NaN  40.821647 -73.950873
       364556                 NaN  40.886361 -73.853290
```

```
364557                 NaN  40.674212 -73.803585

                                      Location    time_elapsed  \
364553   (40.69514470265117, -73.86094888534394) 0 days 10:17:47
364554   (40.86782963689454, -73.90717786644662) 0 days 02:20:34
364555  (40.821646626438095, -73.95087342885292) 0 days 00:19:03
364556   (40.88636077906953, -73.85329048666742) 0 days 02:40:53
364557  (40.674211762243935, -73.80358548685278) 0 days 02:47:00

       time_elapsed_seconds
364553              37067.0
364554               8434.0
364555               1143.0
364556               9653.0
364557              10020.0

[5 rows x 55 columns]
```

Let us find the invalid dates in the column created date, due date ,resolution action updated date,closed date. then remove the invalid entries.

[74]:
```
invalid_dates_count_created_date = service_request['Created Date'].isnull().
 ↪sum()
print(f"Number of invalid dates for creted dates:␣
 ↪{invalid_dates_count_created_date}")
```

```
Number of invalid dates for creted dates: 0
```

[75]:
```
invalid_dates_count_due_date = service_request['Due Date'].isnull().sum()
print(f"Number of invalid dates for Due dates: {invalid_dates_count_due_date}")
```

```
Number of invalid dates for Due dates: 0
```

[76]:
```
invalid_dates_count_Updated_date = service_request['Resolution Action Updated␣
 ↪Date'].isnull().sum()
print(f"Number of invalid dates for ResolutionAction Updted dates :␣
 ↪{invalid_dates_count_Updated_date}")
```

```
Number of invalid dates for ResolutionAction Updted dates : 0
```

[77]:
```
invalid_dates_count_Closed_date = service_request['Closed Date'].isnull().sum()
print(f"Number of invalid dates Closed dates:␣
 ↪{invalid_dates_count_Closed_date}")
```

```
Number of invalid dates Closed dates: 0
```

[78]:
```
service_request['Resolution Action Updated Date'].isnull().sum()
```

[78]: 0

[79]:
```
#Remove the entries that have incorrect timelines

service_request = service_request.dropna(subset=['Resolution Action Updated␣
↪Date','Due Date'],axis=0)
```

[80]:
```
service_request['Resolution Action Updated Date'].isnull().sum()
```

[80]: 0

[81]:
```
service_request['Due Date'].isnull().sum()
```

[81]: 0

[82]:
```
#Calculate the time elapsed in closed and creation date.
#Convert the calculated date to seconds to get a better representation.

service_request['time_elapsed'] = service_request['Closed Date'] -␣
↪service_request['Created Date']

service_request.shape
service_request.head()
```

[82]:
```
   Unique Key        Created Date         Closed Date Agency  \
0    32310363 2015-12-31 23:59:45 2016-01-01 00:55:15   NYPD
1    32309934 2015-12-31 23:59:44 2016-01-01 01:26:57   NYPD
2    32309159 2015-12-31 23:59:29 2016-01-01 04:51:03   NYPD
3    32305098 2015-12-31 23:57:46 2016-01-01 07:43:13   NYPD
4    32306529 2015-12-31 23:56:58 2016-01-01 03:24:42   NYPD

                       Agency Name           Complaint Type  \
0  New York City Police Department  Noise - Street/Sidewalk
1  New York City Police Department         Blocked Driveway
2  New York City Police Department         Blocked Driveway
3  New York City Police Department          Illegal Parking
4  New York City Police Department          Illegal Parking

                     Descriptor    Location Type  Incident Zip  \
0              Loud Music/Party  Street/Sidewalk       10034.0
1                     No Access  Street/Sidewalk       11105.0
2                     No Access  Street/Sidewalk       10458.0
3  Commercial Overnight Parking  Street/Sidewalk       10461.0
4              Blocked Sidewalk  Street/Sidewalk       11373.0

        Incident Address  … Road Ramp Bridge Highway Segment  \
0    71 VERMILYEA AVENUE  …       NaN                    NaN
```

```
1        27-07 23 AVENUE  …       NaN                    NaN
2  2897 VALENTINE AVENUE  …       NaN                    NaN
3    2940 BAISLEY AVENUE  …       NaN                    NaN
4          87-14 57 ROAD  …       NaN                    NaN

  Garage Lot Name Ferry Direction Ferry Terminal Name   Latitude  Longitude  \
0             NaN             NaN                 NaN  40.865682 -73.923501
1             NaN             NaN                 NaN  40.775945 -73.915094
2             NaN             NaN                 NaN  40.870325 -73.888525
3             NaN             NaN                 NaN  40.835994 -73.828379
4             NaN             NaN                 NaN  40.733060 -73.874170

                                   Location    time_elapsed  \
0   (40.86568153633767, -73.92350095571744) 0 days 00:55:30
1  (40.775945312321085, -73.91509393898605) 0 days 01:27:13
2  (40.870324522111424, -73.88852464418646) 0 days 04:51:34
3   (40.83599404683083, -73.82837939584206) 0 days 07:45:27
4  (40.733059618956815, -73.87416975810375) 0 days 03:27:44

  time_elapsed_seconds
0               3330.0
1               5233.0
2              17494.0
3              27927.0
4              12464.0

[5 rows x 55 columns]
```

In the above time elapse hs been calculated by finding the difference in the "close date" and "created date" column and added a new column "time elapsed"

```
[83]: # Convert the time_elapsed to seconds for better representation
service_request['time_elapsed_seconds'] = service_request['time_elapsed'].dt.
 ↪total_seconds()

print(service_request[['Closed Date', 'Created Date', 'time_elapsed',␣
 ↪'time_elapsed_seconds']])
```

```
               Closed Date        Created Date    time_elapsed  \
0      2016-01-01 00:55:15 2015-12-31 23:59:45 0 days 00:55:30
1      2016-01-01 01:26:57 2015-12-31 23:59:44 0 days 01:27:13
2      2016-01-01 04:51:03 2015-12-31 23:59:29 0 days 04:51:34
3      2016-01-01 07:43:13 2015-12-31 23:57:46 0 days 07:45:27
4      2016-01-01 03:24:42 2015-12-31 23:56:58 0 days 03:27:44
…                        …                 …             …
364553 2015-01-01 10:22:31 2015-01-01 00:04:44 0 days 10:17:47
364554 2015-01-01 02:25:02 2015-01-01 00:04:28 0 days 02:20:34
364555 2015-01-01 00:20:33 2015-01-01 00:01:30 0 days 00:19:03
```

```
364556 2015-01-01 02:42:22 2015-01-01 00:01:29 0 days 02:40:53
364557 2015-01-01 02:47:50 2015-01-01 00:00:50 0 days 02:47:00

        time_elapsed_seconds
0                     3330.0
1                     5233.0
2                    17494.0
3                    27927.0
4                    12464.0
…                          …
364553               37067.0
364554                8434.0
364555                1143.0
364556                9653.0
364557               10020.0

[362138 rows x 4 columns]
```

[84]:
```
service_request.shape
```

[84]:
```
(362138, 55)
```

[85]:
```
service_request.dtypes
```

[85]:
```
Unique Key                                 int64
Created Date                      datetime64[ns]
Closed Date                       datetime64[ns]
Agency                                    object
Agency Name                               object
Complaint Type                            object
Descriptor                                object
Location Type                             object
Incident Zip                             float64
Incident Address                          object
Street Name                               object
Cross Street 1                            object
Cross Street 2                            object
Intersection Street 1                     object
Intersection Street 2                     object
Address Type                              object
City                                      object
Landmark                                  object
Facility Type                             object
Status                                    object
Due Date                          datetime64[ns]
Resolution Description                    object
Resolution Action Updated Date    datetime64[ns]
```

```
Community Board                            object
Borough                                    object
X Coordinate (State Plane)                float64
Y Coordinate (State Plane)                float64
Park Facility Name                         object
Park Borough                               object
School Name                                object
School Number                              object
School Region                              object
School Code                                object
School Phone Number                        object
School Address                             object
School City                                object
School State                               object
School Zip                                 object
School Not Found                           object
School or Citywide Complaint              float64
Vehicle Type                              float64
Taxi Company Borough                      float64
Taxi Pick Up Location                     float64
Bridge Highway Name                        object
Bridge Highway Direction                   object
Road Ramp                                  object
Bridge Highway Segment                     object
Garage Lot Name                           float64
Ferry Direction                            object
Ferry Terminal Name                        object
Latitude                                  float64
Longitude                                 float64
Location                                   object
time_elapsed                      timedelta64[ns]
time_elapsed_seconds                      float64
dtype: object
```

0.5 View the descriptive statistics for the newly created columns.i.e. time_elapsed, time_elapsed_seconds.

[36]:

```
print(service_request[['time_elapsed','time_elapsed_seconds']].describe())
```

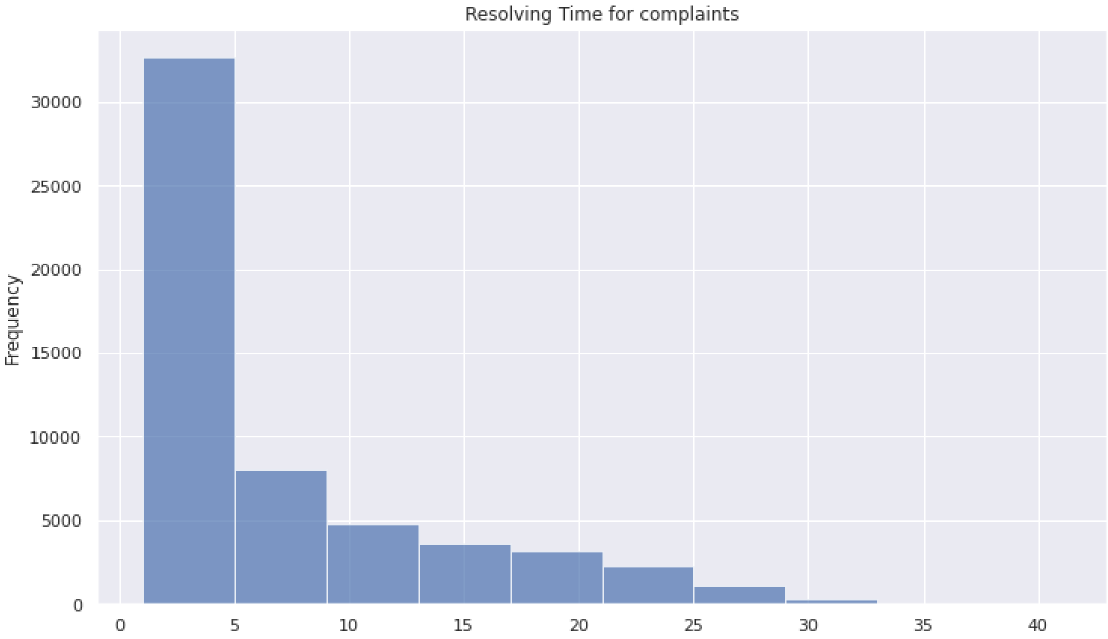
```
                    time_elapsed  time_elapsed_seconds
count                     362138          3.621380e+05
mean   0 days 04:11:53.463779001          1.511346e+04
std    0 days 05:51:43.422780424          2.110342e+04
min              0 days 00:01:01          6.100000e+01
25%              0 days 01:15:33          4.533000e+03
50%              0 days 02:40:16          9.616000e+03
```

```
75%              0 days 05:14:37          1.887700e+04
max             24 days 16:52:22          2.134342e+06
```

Let us see grphically the time_elapsed and time_elapsed_second

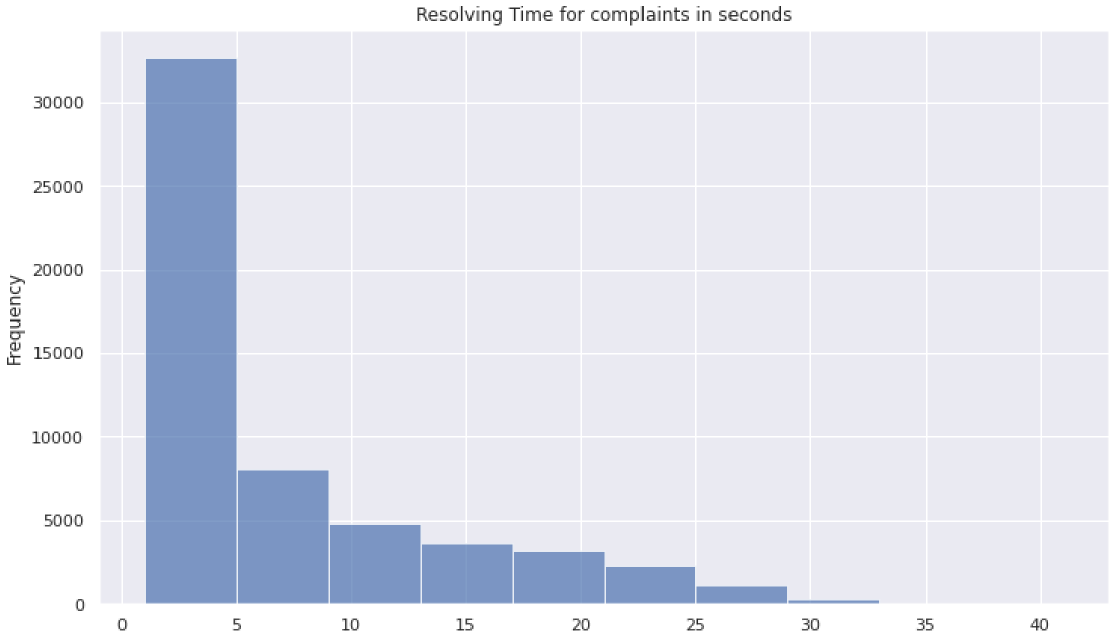
[37]:

```
# descriptive statistics for resolving time from closed and create␣
 ↪date(time_elapsed column)
plt.figure(figsize=(12,7))
service_request['time_elapsed'].value_counts().
 ↪plot(kind='hist',x='time_elapsed',alpha=0.7,title="Resolving Time for␣
 ↪complaints")
plt.show()
```



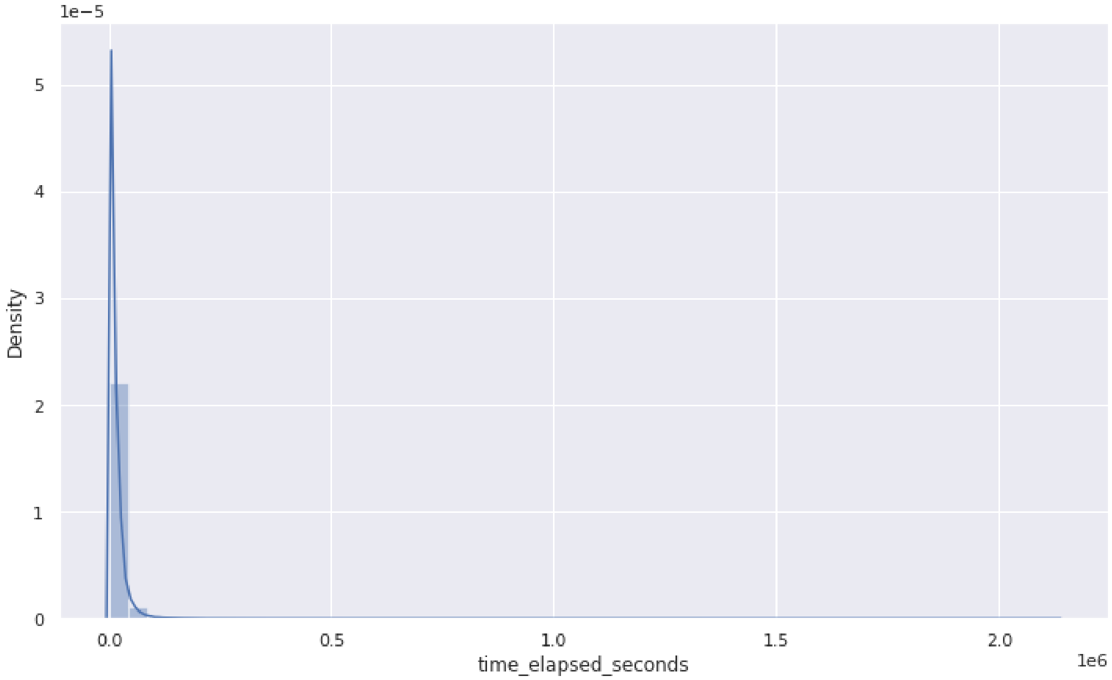
[38]:

```
# descriptive statistics for resolving time from closed and create␣
 ↪date(time_elapsed_seconds column)
plt.figure(figsize=(12,7))
service_request['time_elapsed_seconds'].value_counts().plot(kind='hist',␣
 ↪x='time_elapsed_seconds',alpha=0.7,title="Resolving Time for complaints in␣
 ↪seconds")
plt.show()
```



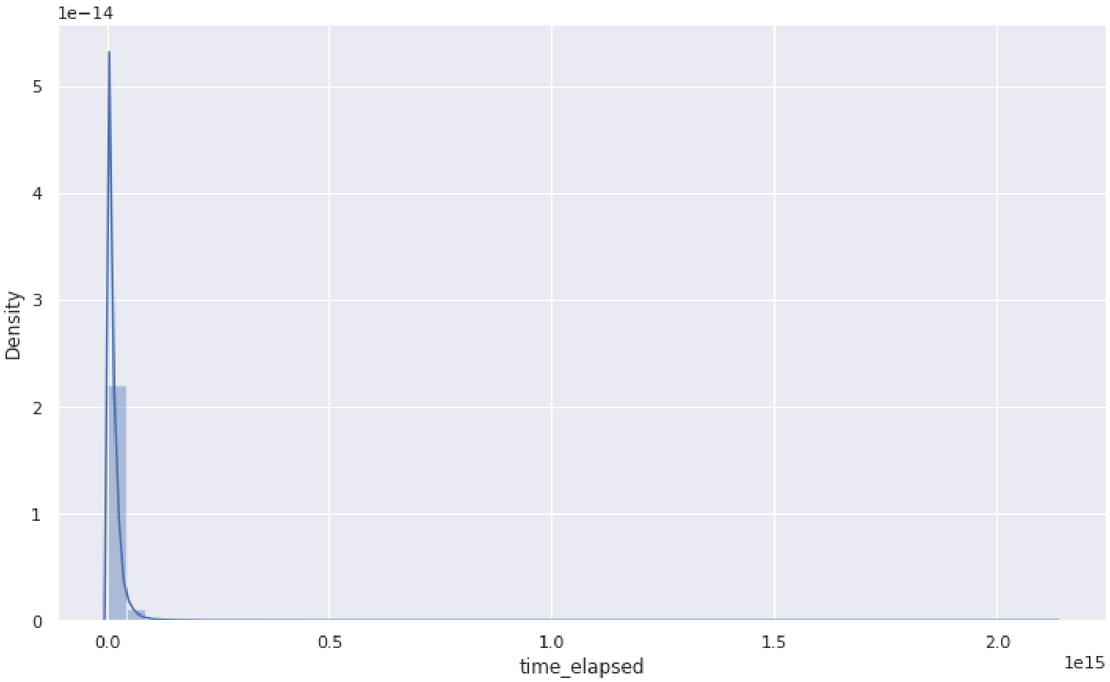
```
[39]: # Alternatively  resolving time for complaints type
      sns.distplot(service_request["time_elapsed_seconds"])

      plt.show
```

```
[39]: <function matplotlib.pyplot.show(close=None, block=None)>
```



[40]:

```
sns.distplot(service_request["time_elapsed"])
plt.show()
```



[41]:

```
service_request.head()
```

[41]:

```
   Unique Key        Created Date         Closed Date Agency  \
0    32310363 2015-12-31 23:59:45 2016-01-01 00:55:15   NYPD
1    32309934 2015-12-31 23:59:44 2016-01-01 01:26:57   NYPD
2    32309159 2015-12-31 23:59:29 2016-01-01 04:51:03   NYPD
3    32305098 2015-12-31 23:57:46 2016-01-01 07:43:13   NYPD
4    32306529 2015-12-31 23:56:58 2016-01-01 03:24:42   NYPD

                       Agency Name           Complaint Type  \
0  New York City Police Department  Noise - Street/Sidewalk
1  New York City Police Department         Blocked Driveway
2  New York City Police Department         Blocked Driveway
3  New York City Police Department          Illegal Parking
4  New York City Police Department          Illegal Parking

                    Descriptor    Location Type  Incident Zip  \
0             Loud Music/Party  Street/Sidewalk       10034.0
1                    No Access  Street/Sidewalk       11105.0
```

```
2                     No Access  Street/Sidewalk       10458.0
3  Commercial Overnight Parking  Street/Sidewalk       10461.0
4               Blocked Sidewalk  Street/Sidewalk       11373.0

        Incident Address  … Road Ramp Bridge Highway Segment  \
0    71 VERMILYEA AVENUE  …       NaN                     NaN
1        27-07 23 AVENUE  …       NaN                     NaN
2  2897 VALENTINE AVENUE  …       NaN                     NaN
3    2940 BAISLEY AVENUE  …       NaN                     NaN
4          87-14 57 ROAD  …       NaN                     NaN

  Garage Lot Name Ferry Direction Ferry Terminal Name   Latitude  Longitude  \
0             NaN             NaN                 NaN  40.865682 -73.923501
1             NaN             NaN                 NaN  40.775945 -73.915094
2             NaN             NaN                 NaN  40.870325 -73.888525
3             NaN             NaN                 NaN  40.835994 -73.828379
4             NaN             NaN                 NaN  40.733060 -73.874170

                                   Location    time_elapsed  \
0   (40.86568153633767, -73.92350095571744) 0 days 00:55:30
1  (40.775945312321085, -73.91509393898605) 0 days 01:27:13
2  (40.870324522111424, -73.88852464418646) 0 days 04:51:34
3   (40.83599404683083, -73.82837939584206) 0 days 07:45:27
4  (40.733059618956815, -73.87416975810375) 0 days 03:27:44

  time_elapsed_seconds
0               3330.0
1               5233.0
2              17494.0
3              27927.0
4              12464.0

[5 rows x 55 columns]
```

0.6 Check the number of null values in the complaint_type and city columns

[42]:

```
null_complaint_type = service_request['Complaint Type'].isnull().sum()
null_city = service_request['City'].isnull().sum()

print(f"Number of null values in 'Complaint_Type': {null_complaint_type}")
print(f"Number of null values in 'City': {null_city}")
```

```
Number of null values in 'Complaint_Type': 0
Number of null values in 'City': 674
```

```
[43]: # Impute the NA value in the 'City' column with 'Unknown City'
      service_request['City'].fillna('Unknown City', inplace=True)
```

```
[44]: service_request["City"].unique()
```

```
[44]: array(['NEW YORK', 'ASTORIA', 'BRONX', 'ELMHURST', 'BROOKLYN',
             'KEW GARDENS', 'JACKSON HEIGHTS', 'MIDDLE VILLAGE', 'REGO PARK',
             'SAINT ALBANS', 'JAMAICA', 'SOUTH RICHMOND HILL', 'Unknown City',
             'RIDGEWOOD', 'HOWARD BEACH', 'FOREST HILLS', 'STATEN ISLAND',
             'OZONE PARK', 'RICHMOND HILL', 'WOODHAVEN', 'FLUSHING', 'CORONA',
             'QUEENS VILLAGE', 'OAKLAND GARDENS', 'HOLLIS', 'MASPETH',
             'EAST ELMHURST', 'SOUTH OZONE PARK', 'WOODSIDE', 'FRESH MEADOWS',
             'LONG ISLAND CITY', 'ROCKAWAY PARK', 'SPRINGFIELD GARDENS',
             'COLLEGE POINT', 'BAYSIDE', 'GLEN OAKS', 'FAR ROCKAWAY',
             'BELLEROSE', 'LITTLE NECK', 'CAMBRIA HEIGHTS', 'ROSEDALE',
             'SUNNYSIDE', 'WHITESTONE', 'ARVERNE', 'FLORAL PARK',
             'NEW HYDE PARK', 'CENTRAL PARK', 'BREEZY POINT', 'QUEENS',
             'Astoria', 'Long Island City', 'Woodside', 'East Elmhurst',
             'Howard Beach'], dtype=object)
```

```
[45]: service_request["City"].nunique()
```

```
[45]: 54
```

```
[46]: # Optionally, you can verify the imputation by checking the null values again
      imputed_null_city = service_request['City'].isnull().sum()
      print(f"Number of null values after imputation: {imputed_null_city}")
```

```
Number of null values after imputation: 0
```

```
[47]: service_request.isna().sum(axis=0)
```

```
[47]: Unique Key                      0
      Created Date                    0
      Closed Date                     0
      Agency                          0
      Agency Name                     0
      Complaint Type                  0
      Descriptor                   6494
      Location Type                  93
      Incident Zip                  675
      Incident Address            51679
      Street Name                 51679
      Cross Street 1              55324
      Cross Street 2              55457
      Intersection Street 1      311517
      Intersection Street 2      311641
```

```
Address Type                         929
City                                   0
Landmark                          361763
Facility Type                         17
Status                                 0
Due Date                               0
Resolution Description                 0
Resolution Action Updated Date         0
Community Board                        0
Borough                                0
X Coordinate (State Plane)          1707
Y Coordinate (State Plane)          1707
Park Facility Name                     0
Park Borough                           0
School Name                            0
School Number                          0
School Region                          1
School Code                            1
School Phone Number                    0
School Address                         0
School City                            0
School State                           0
School Zip                             1
School Not Found                       0
School or Citywide Complaint      362138
Vehicle Type                      362138
Taxi Company Borough              362138
Taxi Pick Up Location             362138
Bridge Highway Name               361841
Bridge Highway Direction          361841
Road Ramp                         361876
Bridge Highway Segment            361876
Garage Lot Name                   362138
Ferry Direction                   362138
Ferry Terminal Name               362138
Latitude                            1707
Longitude                           1707
Location                            1707
time_elapsed                           0
time_elapsed_seconds                   0
dtype: int64
```

```
[48]: service_request.dtypes
```

```
[48]: Unique Key                                  int64
      Created Date                       datetime64[ns]
      Closed Date                        datetime64[ns]
```

```
Agency                                     object
Agency Name                                object
Complaint Type                             object
Descriptor                                 object
Location Type                              object
Incident Zip                              float64
Incident Address                           object
Street Name                                object
Cross Street 1                             object
Cross Street 2                             object
Intersection Street 1                      object
Intersection Street 2                      object
Address Type                               object
City                                       object
Landmark                                   object
Facility Type                              object
Status                                     object
Due Date                           datetime64[ns]
Resolution Description                     object
Resolution Action Updated Date     datetime64[ns]
Community Board                            object
Borough                                    object
X Coordinate (State Plane)                float64
Y Coordinate (State Plane)                float64
Park Facility Name                         object
Park Borough                               object
School Name                                object
School Number                              object
School Region                              object
School Code                                object
School Phone Number                        object
School Address                             object
School City                                object
School State                               object
School Zip                                 object
School Not Found                           object
School or Citywide Complaint              float64
Vehicle Type                              float64
Taxi Company Borough                      float64
Taxi Pick Up Location                     float64
Bridge Highway Name                        object
Bridge Highway Direction                   object
Road Ramp                                  object
Bridge Highway Segment                     object
Garage Lot Name                           float64
Ferry Direction                            object
Ferry Terminal Name                        object
```

```
Latitude                                  float64
Longitude                                 float64
Location                                   object
time_elapsed                      timedelta64[ns]
time_elapsed_seconds                      float64
dtype: object
```

[49]:

```
service_request[['Complaint Type','City']]
```

[49]:

```
                 Complaint Type               City
0       Noise - Street/Sidewalk           NEW YORK
1              Blocked Driveway            ASTORIA
2              Blocked Driveway              BRONX
3               Illegal Parking              BRONX
4               Illegal Parking          ELMHURST
…                           …                  …
364553          Illegal Parking          WOODHAVEN
364554          Noise - Vehicle              BRONX
364555  Noise - Street/Sidewalk           NEW YORK
364556         Blocked Driveway              BRONX
364557         Blocked Driveway  SOUTH OZONE PARK

[362138 rows x 2 columns]
```

[50]:

```
Complaint_type= service_request['Complaint Type'].value_counts()
City_type= service_request['City'].value_counts()
print('Complaint types in the city:',Complaint_type)
print('*****'*40)
print('City that has complaints:',City_type)
```

```
Complaint types in the city: Blocked Driveway             100618
Illegal Parking              91705
Noise - Street/Sidewalk      51131
Noise - Commercial           43749
Derelict Vehicle             21516
Noise - Vehicle              19300
Animal Abuse                 10530
Traffic                       5193
Homeless Encampment           4877
Vending                       4183
Noise - Park                  4088
Drinking                      1404
Noise - House of Worship      1068
Posting Advertisement          678
Urinating in Public            641
Bike/Roller/Skate Chronic      475
Panhandling                    325
```

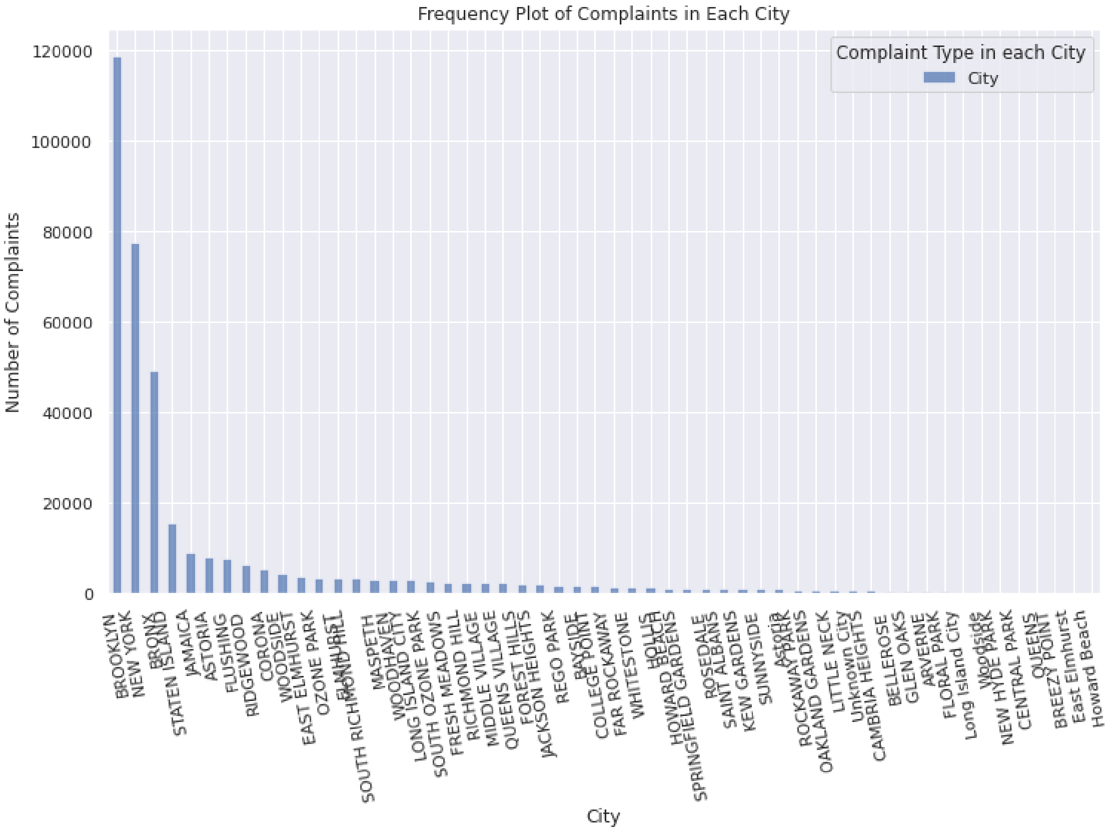
```
Disorderly Youth                315
Illegal Fireworks               172
Graffiti                        157
Agency Issues                     8
Squeegee                          4
Animal in a Park                  1
Name: Complaint Type, dtype: int64
********************************************************************************
********************************************************************************
****************************************
City that has complaints: BROOKLYN               118841
NEW YORK                 77281
BRONX                    49163
STATEN ISLAND            15332
JAMAICA                   8921
ASTORIA                   7991
FLUSHING                  7486
RIDGEWOOD                 6391
CORONA                    5382
WOODSIDE                  4356
EAST ELMHURST             3557
OZONE PARK                3446
ELMHURST                  3437
SOUTH RICHMOND HILL       3431
MASPETH                   3117
WOODHAVEN                 3101
LONG ISLAND CITY          3026
SOUTH OZONE PARK          2668
FRESH MEADOWS             2451
RICHMOND HILL             2333
MIDDLE VILLAGE            2291
QUEENS VILLAGE            2251
FOREST HILLS              2122
JACKSON HEIGHTS           2105
REGO PARK                 1807
BAYSIDE                   1550
COLLEGE POINT             1544
FAR ROCKAWAY              1397
WHITESTONE                1369
HOLLIS                    1231
HOWARD BEACH              1144
SPRINGFIELD GARDENS       1093
ROSEDALE                  1091
SAINT ALBANS              1047
KEW GARDENS               1008
SUNNYSIDE                  944
Astoria                    905
ROCKAWAY PARK              831
```

```
OAKLAND GARDENS          717
LITTLE NECK              712
Unknown City             674
CAMBRIA HEIGHTS          617
BELLEROSE                487
GLEN OAKS                361
ARVERNE                  259
FLORAL PARK              196
Long Island City         170
Woodside                 166
NEW HYDE PARK            129
CENTRAL PARK             110
QUEENS                    37
BREEZY POINT              31
East Elmhurst             30
Howard Beach               1
Name: City, dtype: int64
```

0.7 Draw a frequency plot for the complaints in each city

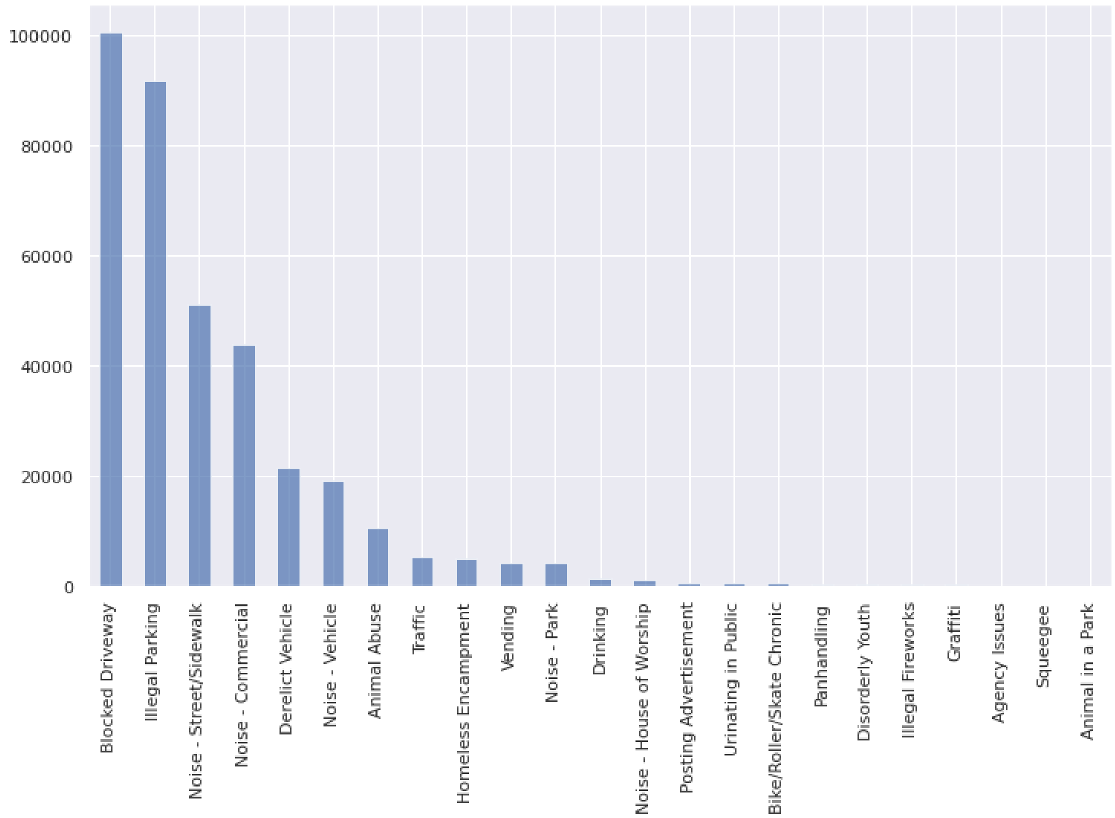
[51]:

```python
service_request['City'].value_counts().plot(kind='bar',alpha=0.7,linewidth=0.8)
plt.xticks(rotation=100)  # Rotate x-axis labels for better readability
plt.xlabel('City')
plt.ylabel('Number of Complaints')
plt.title('Frequency Plot of Complaints in Each City')
plt.legend(title='Complaint Type in each City')
plt.show()
```



[52]:

```
# Let us understand the complaint type via bar graph:
plt.figure(figsize=(12,7))
service_request['Complaint Type'].value_counts().plot(kind='bar',alpha=0.7)
plt.show()
```



[53]:

```
Complaint_type= service_request['Complaint Type'].value_counts()
City_type =service_request['City'].value_counts()
print(Complaint_type)
print('******'* 40)
print(City_type)
```

```
Blocked Driveway            100618
Illegal Parking              91705
Noise - Street/Sidewalk      51131
Noise - Commercial           43749
Derelict Vehicle             21516
Noise - Vehicle              19300
Animal Abuse                 10530
Traffic                       5193
Homeless Encampment           4877
Vending                       4183
Noise - Park                  4088
Drinking                      1404
Noise - House of Worship      1068
Posting Advertisement          678
Urinating in Public            641
```

```
Bike/Roller/Skate Chronic       475
Panhandling                     325
Disorderly Youth                315
Illegal Fireworks               172
Graffiti                        157
Agency Issues                     8
Squeegee                          4
Animal in a Park                  1
Name: Complaint Type, dtype: int64
********************************************************************************
********************************************************************************
********************************************************************************
BROOKLYN               118841
NEW YORK                77281
BRONX                   49163
STATEN ISLAND           15332
JAMAICA                  8921
ASTORIA                  7991
FLUSHING                 7486
RIDGEWOOD                6391
CORONA                   5382
WOODSIDE                 4356
EAST ELMHURST            3557
OZONE PARK               3446
ELMHURST                 3437
SOUTH RICHMOND HILL      3431
MASPETH                  3117
WOODHAVEN                3101
LONG ISLAND CITY         3026
SOUTH OZONE PARK         2668
FRESH MEADOWS            2451
RICHMOND HILL            2333
MIDDLE VILLAGE           2291
QUEENS VILLAGE           2251
FOREST HILLS             2122
JACKSON HEIGHTS          2105
REGO PARK                1807
BAYSIDE                  1550
COLLEGE POINT            1544
FAR ROCKAWAY             1397
WHITESTONE               1369
HOLLIS                   1231
HOWARD BEACH             1144
SPRINGFIELD GARDENS      1093
ROSEDALE                 1091
SAINT ALBANS             1047
KEW GARDENS              1008
SUNNYSIDE                 944
```

```
Astoria                      905
ROCKAWAY PARK                831
OAKLAND GARDENS              717
LITTLE NECK                  712
Unknown City                 674
CAMBRIA HEIGHTS              617
BELLEROSE                    487
GLEN OAKS                    361
ARVERNE                      259
FLORAL PARK                  196
Long Island City             170
Woodside                     166
NEW HYDE PARK                129
CENTRAL PARK                 110
QUEENS                        37
BREEZY POINT                  31
East Elmhurst                 30
Howard Beach                   1
Name: City, dtype: int64
```

0.8 Create Scatter and hexbin plot of the concentration of complaints across Brooklyn

[54]:

```python
# selecting rows based on condition
Brooklyn_city= service_request.loc[service_request['City']=="BROOKLYN"]

print('\nResult dataframe :\n', Brooklyn_city)
```

```
Result dataframe :
         Unique Key        Created Date         Closed Date Agency  \
5         32306554 2015-12-31 23:56:30 2016-01-01 01:50:11   NYPD
9         32308391 2015-12-31 23:53:58 2016-01-01 01:17:40   NYPD
13        32305074 2015-12-31 23:47:58 2016-01-01 08:18:47   NYPD
17        32310273 2015-12-31 23:44:52 2016-01-01 00:36:10   NYPD
18        32306617 2015-12-31 23:40:59 2016-01-01 02:37:28   NYPD
…            …                 …                 …    …
364539    29608505 2015-01-01 00:23:55 2015-01-01 02:58:38   NYPD
364541    29612697 2015-01-01 00:19:22 2015-01-01 02:41:10   NYPD
364544    29613295 2015-01-01 00:17:48 2015-01-01 03:24:48   NYPD
364545    29613456 2015-01-01 00:17:47 2015-01-01 00:51:13   NYPD
364546    29613402 2015-01-01 00:15:45 2015-01-01 02:04:54   NYPD

                            Agency Name      Complaint Type  \
5       New York City Police Department     Illegal Parking
9       New York City Police Department    Blocked Driveway
```

```
13      New York City Police Department     Illegal Parking
17      New York City Police Department  Noise - Commercial
18      New York City Police Department  Noise - Commercial
…                                 …                 …
364539  New York City Police Department    Blocked Driveway
364541  New York City Police Department    Blocked Driveway
364544  New York City Police Department  Noise - Commercial
364545  New York City Police Department    Blocked Driveway
364546  New York City Police Department    Blocked Driveway

                           Descriptor        Location Type  Incident Zip  \
5       Posted Parking Sign Violation      Street/Sidewalk       11215.0
9                           No Access      Street/Sidewalk       11219.0
13      Posted Parking Sign Violation      Street/Sidewalk       11208.0
17                   Loud Music/Party  Club/Bar/Restaurant       11217.0
18                   Loud Music/Party  Club/Bar/Restaurant       11234.0
…                                 …                   …            …
364539                      No Access      Street/Sidewalk       11201.0
364541                      No Access      Street/Sidewalk       11211.0
364544               Loud Music/Party     Store/Commercial       11217.0
364545                      No Access      Street/Sidewalk       11237.0
364546                      No Access      Street/Sidewalk       11218.0

            Incident Address  … Road Ramp Bridge Highway Segment  \
5              260 21 STREET  …       NaN                    NaN
9             1408 66 STREET  …       NaN                    NaN
13              38 COX PLACE  …       NaN                    NaN
17         622 DEGRAW STREET  …       NaN                    NaN
18      2192 FLATBUSH AVENUE  …       NaN                    NaN
…                        …  …       …                    …
364539   229 DUFFIELD STREET  …       NaN                    NaN
364541        27 HOPE STREET  …       NaN                    NaN
364544   242 FLATBUSH AVENUE  …       NaN                    NaN
364545   1373 DECATUR STREET  …       NaN                    NaN
364546      19 MICIELI PLACE  …       NaN                    NaN

       Garage Lot Name Ferry Direction Ferry Terminal Name   Latitude  \
5                  NaN             NaN                 NaN  40.660823
9                  NaN             NaN                 NaN  40.623793
13                 NaN             NaN                 NaN  40.687511
17                 NaN             NaN                 NaN  40.679154
18                 NaN             NaN                 NaN  40.616550
…                …             …                 …        …
364539             NaN             NaN                 NaN  40.691248
364541             NaN             NaN                 NaN  40.713613
364544             NaN             NaN                 NaN  40.680352
364545             NaN             NaN                 NaN  40.692384
364546             NaN             NaN                 NaN  40.646439
```

```
        Longitude                                   Location    time_elapsed  \
5      -73.992568   (40.66082272389114, -73.99256786342693) 0 days 01:53:41
9      -73.999539  (40.623793065806524, -73.99953890121567) 0 days 01:23:42
13     -73.874505   (40.68751060232221, -73.87450451131276) 0 days 08:30:49
17     -73.983430  (40.679154133157326, -73.98342992763081) 0 days 00:51:18
18     -73.930202   (40.61655032892211, -73.93020153359745) 0 days 02:56:29
…             …                                         …              …
364539 -73.984375   (40.69124772858873, -73.98437529459297) 0 days 02:34:43
364541 -73.956966   (40.71361332917013, -73.95696623841229) 0 days 02:21:48
364544 -73.974766   (40.68035202521423, -73.97476587888686) 0 days 03:07:00
364545 -73.904011  (40.692383519190834, -73.90401080101479) 0 days 00:33:26
364546 -73.981971   (40.64643889447912, -73.98197140465561) 0 days 01:49:09

       time_elapsed_seconds
5                    6821.0
9                    5022.0
13                  30649.0
17                   3078.0
18                  10589.0
…                       …
364539               9283.0
364541               8508.0
364544              11220.0
364545               2006.0
364546               6549.0

[118841 rows x 55 columns]
```

[55]:
```
service_request.shape
```

[55]:
```
(362138, 55)
```

[56]:
```
service_request.head()
```

[56]:
```
   Unique Key        Created Date         Closed Date Agency  \
0    32310363 2015-12-31 23:59:45 2016-01-01 00:55:15   NYPD
1    32309934 2015-12-31 23:59:44 2016-01-01 01:26:57   NYPD
2    32309159 2015-12-31 23:59:29 2016-01-01 04:51:03   NYPD
3    32305098 2015-12-31 23:57:46 2016-01-01 07:43:13   NYPD
4    32306529 2015-12-31 23:56:58 2016-01-01 03:24:42   NYPD

                       Agency Name           Complaint Type  \
0  New York City Police Department  Noise - Street/Sidewalk
1  New York City Police Department         Blocked Driveway
2  New York City Police Department         Blocked Driveway
3  New York City Police Department           Illegal Parking
```

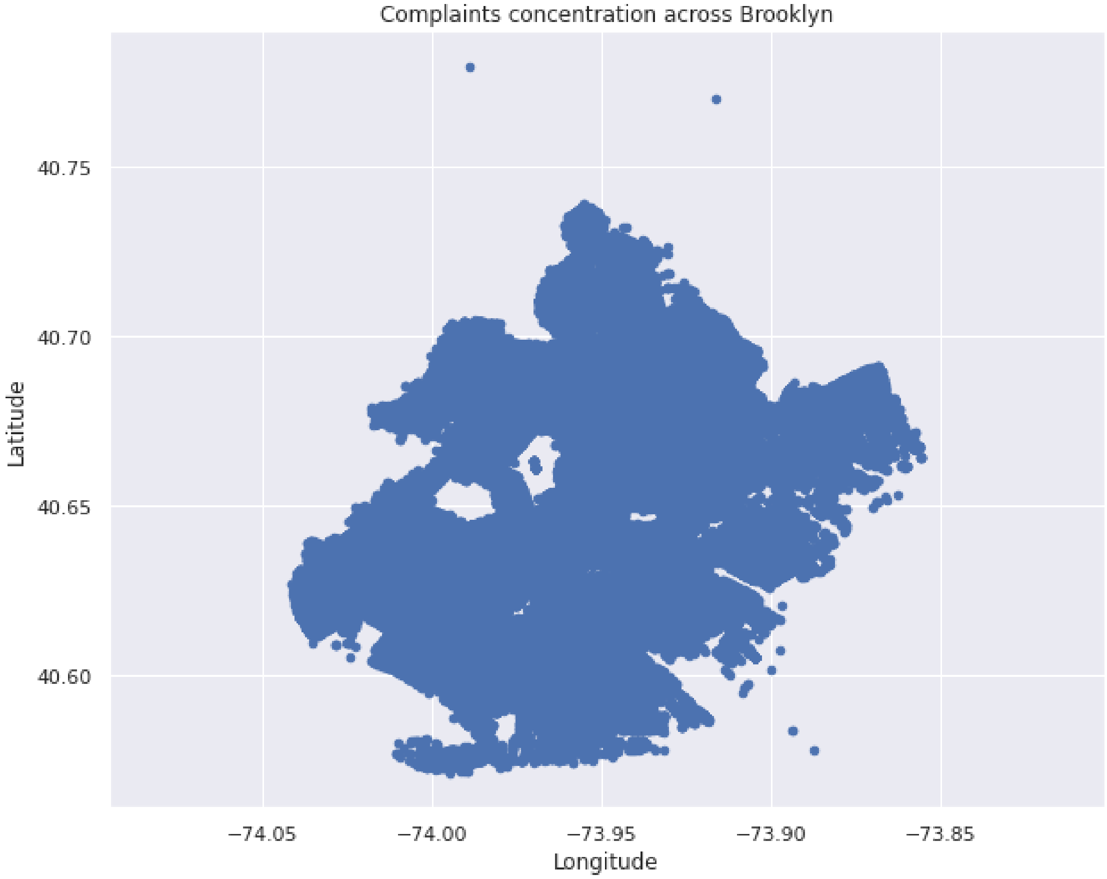
```
4  New York City Police Department          Illegal Parking

                    Descriptor    Location Type  Incident Zip  \
0              Loud Music/Party  Street/Sidewalk       10034.0
1                     No Access  Street/Sidewalk       11105.0
2                     No Access  Street/Sidewalk       10458.0
3  Commercial Overnight Parking  Street/Sidewalk       10461.0
4              Blocked Sidewalk  Street/Sidewalk       11373.0

        Incident Address  … Road Ramp Bridge Highway Segment  \
0    71 VERMILYEA AVENUE  …       NaN                     NaN
1        27-07 23 AVENUE  …       NaN                     NaN
2  2897 VALENTINE AVENUE  …       NaN                     NaN
3    2940 BAISLEY AVENUE  …       NaN                     NaN
4          87-14 57 ROAD  …       NaN                     NaN

  Garage Lot Name Ferry Direction Ferry Terminal Name   Latitude  Longitude  \
0             NaN             NaN                 NaN  40.865682 -73.923501
1             NaN             NaN                 NaN  40.775945 -73.915094
2             NaN             NaN                 NaN  40.870325 -73.888525
3             NaN             NaN                 NaN  40.835994 -73.828379
4             NaN             NaN                 NaN  40.733060 -73.874170

                                  Location    time_elapsed  \
0   (40.86568153633767, -73.92350095571744) 0 days 00:55:30
1  (40.775945312321085, -73.91509393898605) 0 days 01:27:13
2  (40.870324522111424, -73.88852464418646) 0 days 04:51:34
3   (40.83599404683083, -73.82837939584206) 0 days 07:45:27
4  (40.733059618956815, -73.87416975810375) 0 days 03:27:44

  time_elapsed_seconds
0               3330.0
1               5233.0
2              17494.0
3              27927.0
4              12464.0

[5 rows x 55 columns]
```

1 Let us see the complaint type in Brooklyn via Graph

[57]:
```
Brooklyn_city[['Longitude', 'Latitude']].plot(kind='scatter',
    x='Longitude', y='Latitude', figsize=(10,8),title = 'Complaints␣
↪concentration across Brooklyn').axis('equal')
```

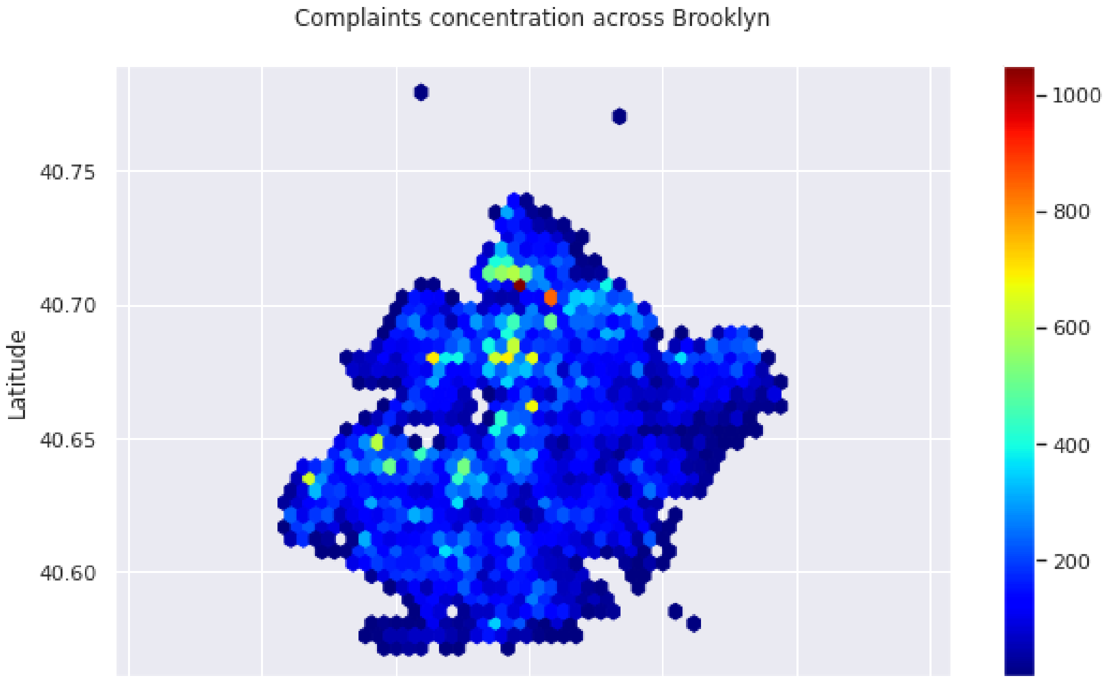
```
*c* argument looks like a single numeric RGB or RGBA sequence, which should be
avoided as value-mapping will have precedence in case its length matches with
*x* & *y*.  Please use the *color* keyword-argument or provide a 2D array with a
single row if you intend to specify the same RGB or RGBA value for all points.
```

[57]:
```
(-74.05061403028367, -73.84647934348564, 40.56112685375488, 40.78979838623255)
```



[58]:
```
Brooklyn_city.plot(kind='hexbin', x='Longitude', y='Latitude', gridsize=40,
    colormap = 'jet',mincnt=1,title = 'Complaints concentration across␣
 ↪Brooklyn\n', figsize=(10,6)).axis('equal')
```

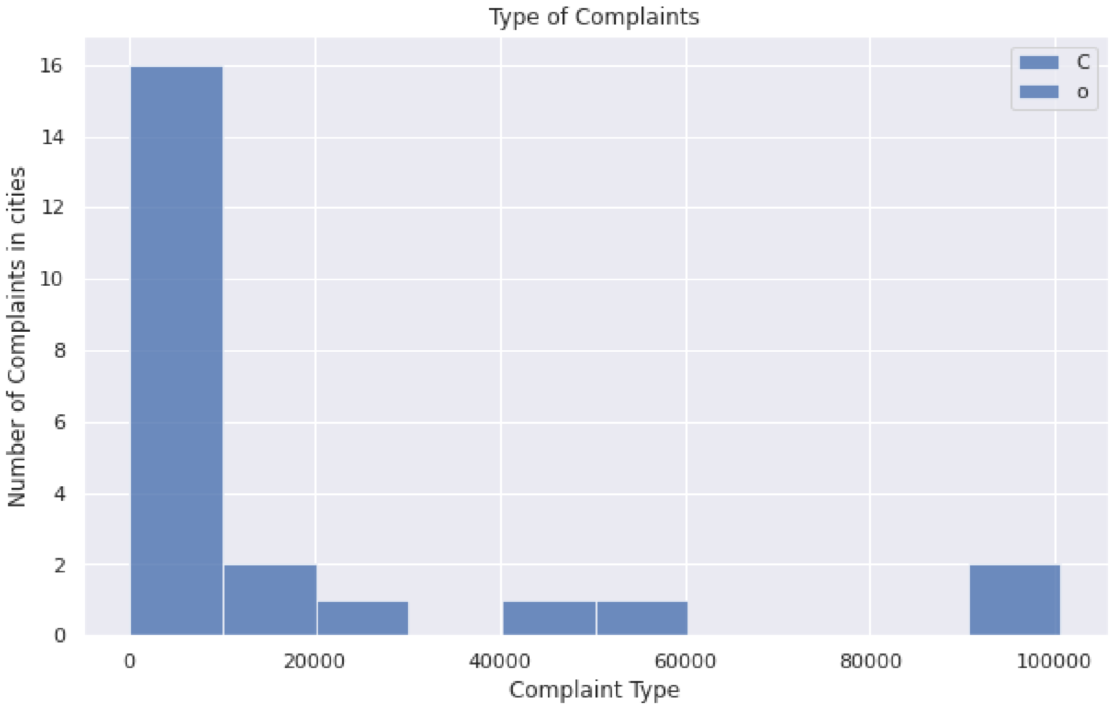
[58]:
```
(-74.05061403048781, -73.8464793432815, 40.56112685375488, 40.78979838623255)
```



1.1 Find the major types of complaints

[59]:

```
# Plot  bar graph to show the types of complaints:
service_request['Complaint Type'].value_counts().plot(kind='hist',
                    figsize=(10,6),alpha=0.8,
                    linewidth=0.8,
                    title = 'Types of Complaints')
plt.xlabel('Complaint Type')
plt.ylabel('Number of Complaints in cities')
plt.title('Type of Complaints')
plt.legend('Complaint type')
plt.show()
```

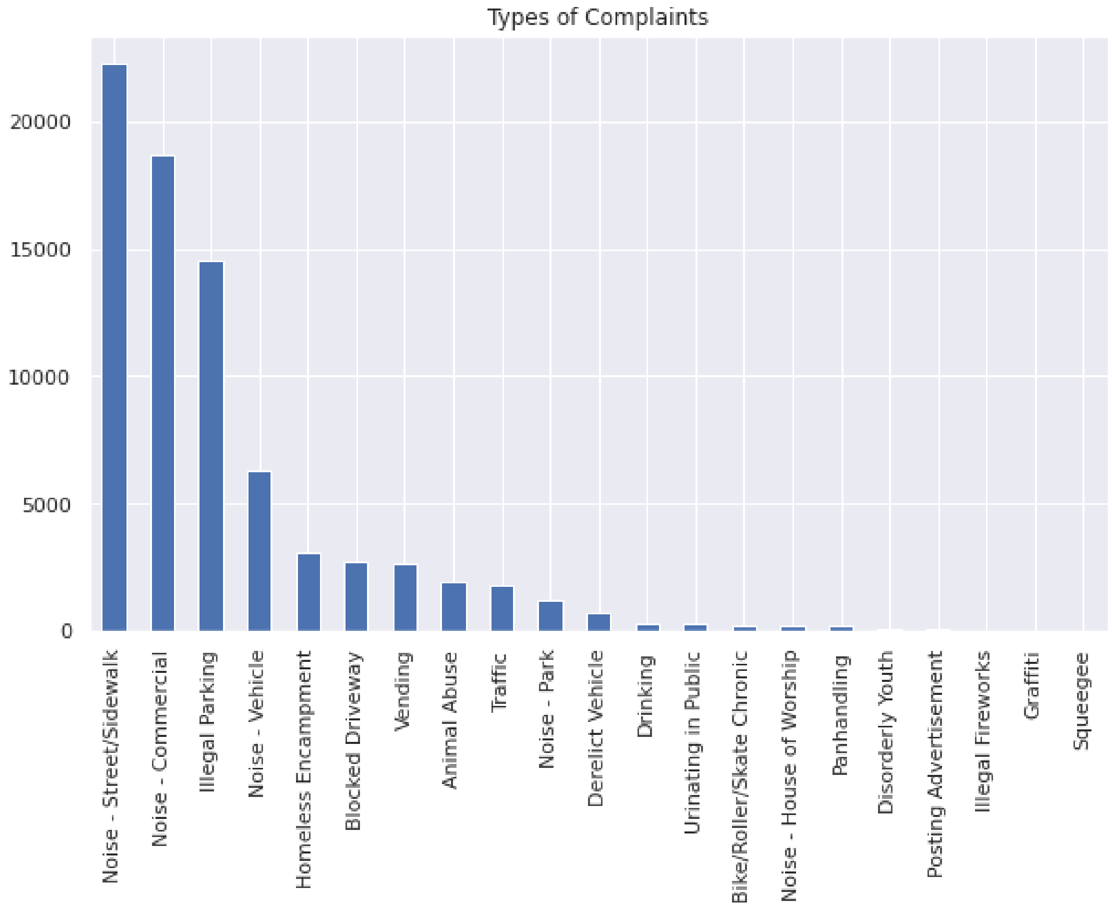


Check the frequency of vrious types of complaints for New York City

```
[60]: New_York_City_data= service_request.loc[service_request['City']=='NEW YORK']
```

```
[61]: New_York_City_data['Complaint Type'].value_counts().plot(kind='bar',
                          figsize=(10,6), title = 'Types of Complaints')
```

```
[61]: <AxesSubplot:title={'center':'Types of Complaints'}>
```



[62]:

```
# find the top ten types of complaints. lets take hundred complaints and look␣
↪for 10 unique complaints.
top_10_complaints = service_request['Complaint Type'].value_counts().
↪nlargest(10)
print("The top 10 Complaint in the cities are:",top_10_complaints)
```

```
The top 10 Complaint in the cities are: Blocked Driveway           100618
Illegal Parking             91705
Noise - Street/Sidewalk     51131
Noise - Commercial          43749
Derelict Vehicle            21516
Noise - Vehicle             19300
Animal Abuse                10530
Traffic                      5193
Homeless Encampment          4877
Vending                      4183
Name: Complaint Type, dtype: int64
```

```
[ ]:
```

1.2 Display the various types of complaints in each city

```
[63]: complaint_by_city =service_request.pivot_table(index='Complaint Type',␣
       ↪columns='City',values='Unique Key',aggfunc='count')
      print(complaint_by_city)
```

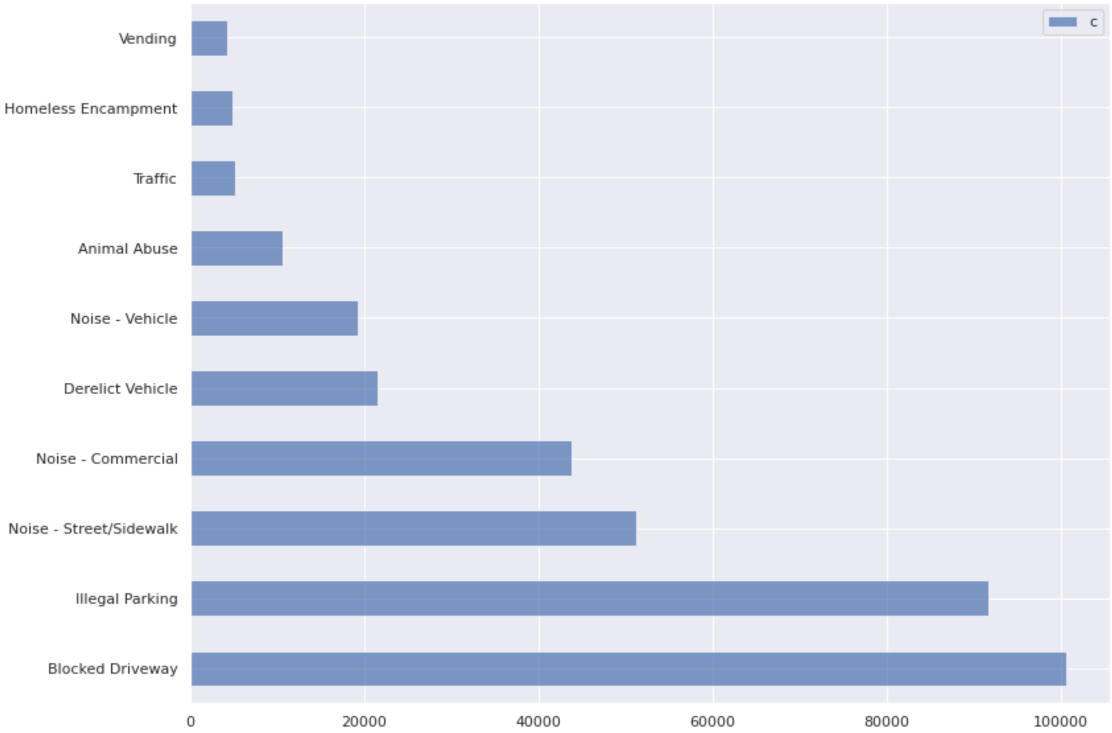
```
City                         ARVERNE  ASTORIA  Astoria  BAYSIDE  BELLEROSE  \
Complaint Type
Agency Issues                    NaN      NaN      NaN      NaN        NaN
Animal Abuse                    46.0    170.0      NaN     53.0       15.0
Animal in a Park                 NaN      NaN      NaN      NaN        NaN
Bike/Roller/Skate Chronic        NaN     16.0      NaN      NaN        1.0
Blocked Driveway                50.0   3436.0    159.0    514.0      138.0
Derelict Vehicle                32.0    426.0     14.0    231.0      120.0
Disorderly Youth                 2.0      5.0      NaN      2.0        2.0
Drinking                         1.0     43.0      NaN      1.0        1.0
Graffiti                         1.0      4.0      NaN      3.0        NaN
Homeless Encampment              4.0     32.0      NaN      2.0        1.0
Illegal Fireworks                NaN      4.0      NaN      NaN        1.0
Illegal Parking                 62.0   1340.0    277.0    638.0      132.0
Noise - Commercial               2.0   1653.0    310.0     47.0       38.0
Noise - House of Worship        14.0     21.0      NaN      3.0        1.0
Noise - Park                     2.0     64.0      NaN      4.0        1.0
Noise - Street/Sidewalk         29.0    409.0    145.0     17.0       13.0
Noise - Vehicle                 10.0    236.0      NaN     24.0       11.0
Panhandling                      1.0      2.0      NaN      NaN        1.0
Posting Advertisement            NaN      3.0      NaN      NaN        1.0
Squeegee                         NaN      NaN      NaN      NaN        NaN
Traffic                          1.0     60.0      NaN      9.0        9.0
Urinating in Public              1.0     10.0      NaN      NaN        1.0
Vending                          1.0     57.0      NaN      2.0        NaN

City                         BREEZY POINT    BRONX  BROOKLYN  CAMBRIA HEIGHTS  \
Complaint Type
Agency Issues                         NaN      NaN       NaN              NaN
Animal Abuse                          2.0   1971.0    3191.0             15.0
Animal in a Park                      NaN      NaN       NaN              NaN
Bike/Roller/Skate Chronic             NaN     22.0     124.0              NaN
Blocked Driveway                      3.0  17061.0   36444.0            177.0
Derelict Vehicle                      3.0   2402.0    6257.0            148.0
Disorderly Youth                      NaN     66.0      79.0              NaN
Drinking                              1.0    206.0     291.0              NaN
Graffiti                              NaN     15.0      60.0              NaN
Homeless Encampment                   NaN    275.0     947.0              6.0
```

```
Illegal Fireworks                   NaN     24.0      61.0              1.0
Illegal Parking                    16.0   9889.0   33528.0            113.0
Noise - Commercial                  4.0   2944.0   13855.0             19.0
Noise - House of Worship            NaN     90.0     389.0              2.0
Noise - Park                        NaN    548.0    1574.0              NaN
Noise - Street/Sidewalk             1.0   9144.0   13982.0             29.0
Noise - Vehicle                     1.0   3555.0    5965.0            100.0
Panhandling                         NaN     20.0      49.0              NaN
Posting Advertisement               NaN     17.0      58.0              NaN
Squeegee                            NaN      NaN       NaN              NaN
Traffic                             NaN    427.0    1258.0              7.0
Urinating in Public                 NaN     54.0     155.0              NaN
Vending                             NaN    433.0     574.0              NaN

City                       CENTRAL PARK  …  SOUTH OZONE PARK  \
Complaint Type                           …
Agency Issues                       NaN  …               NaN
Animal Abuse                        NaN  …              74.0
Animal in a Park                    NaN  …               NaN
Bike/Roller/Skate Chronic           NaN  …               1.0
Blocked Driveway                    NaN  …            1202.0
Derelict Vehicle                    NaN  …             425.0
Disorderly Youth                    NaN  …               2.0
Drinking                            NaN  …              14.0
Graffiti                            NaN  …               2.0
Homeless Encampment                 NaN  …               5.0
Illegal Fireworks                   NaN  …               1.0
Illegal Parking                     5.0  …             602.0
Noise - Commercial                  NaN  …              82.0
Noise - House of Worship            NaN  …               5.0
Noise - Park                        NaN  …               4.0
Noise - Street/Sidewalk           105.0  …             108.0
Noise - Vehicle                     NaN  …              97.0
Panhandling                         NaN  …               NaN
Posting Advertisement               NaN  …               1.0
Squeegee                            NaN  …               NaN
Traffic                             NaN  …              36.0
Urinating in Public                 NaN  …               2.0
Vending                             NaN  …               5.0

City                       SOUTH RICHMOND HILL  SPRINGFIELD GARDENS  \
Complaint Type
Agency Issues                              NaN                  NaN
Animal Abuse                              40.0                 42.0
Animal in a Park                           NaN                  NaN
Bike/Roller/Skate Chronic                  1.0                  NaN
Blocked Driveway                        1946.0                330.0
Derelict Vehicle                         356.0                266.0
```

```
Disorderly Youth                           2.0                  NaN
Drinking                                  25.0                  6.0
Graffiti                                   NaN                  NaN
Homeless Encampment                       12.0                  7.0
Illegal Fireworks                          2.0                  1.0
Illegal Parking                          596.0                291.0
Noise - Commercial                       223.0                 38.0
Noise - House of Worship                   3.0                  1.0
Noise - Park                               2.0                  1.0
Noise - Street/Sidewalk                   93.0                 42.0
Noise - Vehicle                           93.0                 48.0
Panhandling                                NaN                  2.0
Posting Advertisement                      NaN                  2.0
Squeegee                                   NaN                  NaN
Traffic                                   12.0                 12.0
Urinating in Public                        1.0                  3.0
Vending                                   24.0                  1.0

City                       STATEN ISLAND  SUNNYSIDE  Unknown City  WHITESTONE  \
Complaint Type
Agency Issues                        NaN        NaN           8.0         NaN
Animal Abuse                       786.0       40.0           1.0        43.0
Animal in a Park                     NaN        NaN           NaN         NaN
Bike/Roller/Skate Chronic           10.0        2.0           2.0         4.0
Blocked Driveway                  2845.0      278.0          86.0       279.0
Derelict Vehicle                  2183.0       17.0          63.0       279.0
Disorderly Youth                    25.0        2.0           NaN         1.0
Drinking                           188.0       12.0           3.0         3.0
Graffiti                             6.0        1.0           NaN         1.0
Homeless Encampment                 77.0       12.0           1.0         NaN
Illegal Fireworks                   11.0        NaN           NaN         1.0
Illegal Parking                   6223.0      167.0         312.0       631.0
Noise - Commercial                 783.0      238.0          79.0        21.0
Noise - House of Worship            18.0        NaN           NaN         NaN
Noise - Park                        67.0       16.0           7.0         7.0
Noise - Street/Sidewalk            885.0       69.0          99.0        35.0
Noise - Vehicle                    424.0       53.0           9.0        31.0
Panhandling                         13.0        NaN           1.0         NaN
Posting Advertisement              516.0        3.0           NaN         NaN
Squeegee                             NaN        NaN           NaN         NaN
Traffic                            228.0       17.0           2.0        32.0
Urinating in Public                 19.0        2.0           NaN         NaN
Vending                             25.0       15.0           1.0         1.0

City                       WOODHAVEN  WOODSIDE  Woodside
Complaint Type
Agency Issues                    NaN       NaN       NaN
Animal Abuse                    57.0     111.0       NaN
Animal in a Park                 NaN       NaN       NaN
Bike/Roller/Skate Chronic        2.0       5.0       NaN
Blocked Driveway              1363.0    2037.0      27.0
Derelict Vehicle               369.0     298.0       8.0
Disorderly Youth                 NaN       1.0       NaN
Drinking                         4.0      15.0       NaN
Graffiti                         NaN       4.0       NaN
Homeless Encampment             10.0      38.0       NaN
Illegal Fireworks                NaN       1.0       NaN
Illegal Parking                895.0    1083.0     124.0
Noise - Commercial             209.0     256.0       2.0
Noise - House of Worship         3.0       4.0       NaN
Noise - Park                     3.0      38.0       NaN
Noise - Street/Sidewalk         89.0     261.0       5.0
Noise - Vehicle                 81.0     136.0       NaN
Panhandling                      1.0       NaN       NaN
Posting Advertisement            NaN       NaN       NaN
Squeegee                         NaN       NaN       NaN
Traffic                          7.0      45.0       NaN
Urinating in Public              2.0       8.0       NaN
Vending                          6.0      15.0       NaN

[23 rows x 54 columns]
```

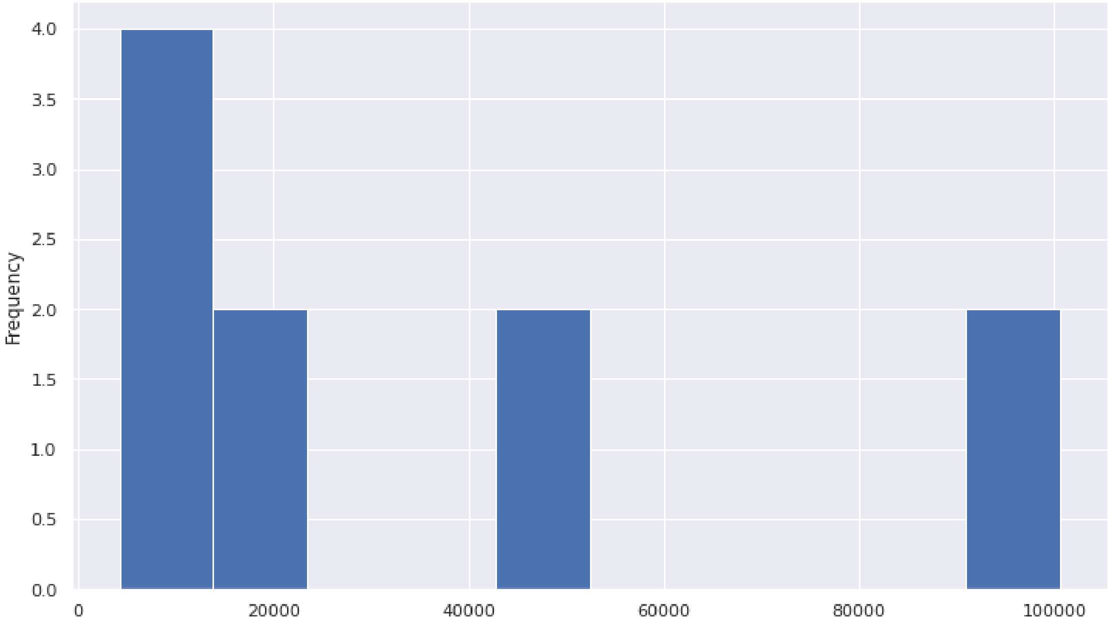
[64]:

```
service_request["Complaint Type"].value_counts()[:10].plot(kind='barh',alpha=0.
 ↪7,figsize=(12,8))
plt.tight_layout()
plt.legend("city")
plt.show()
```

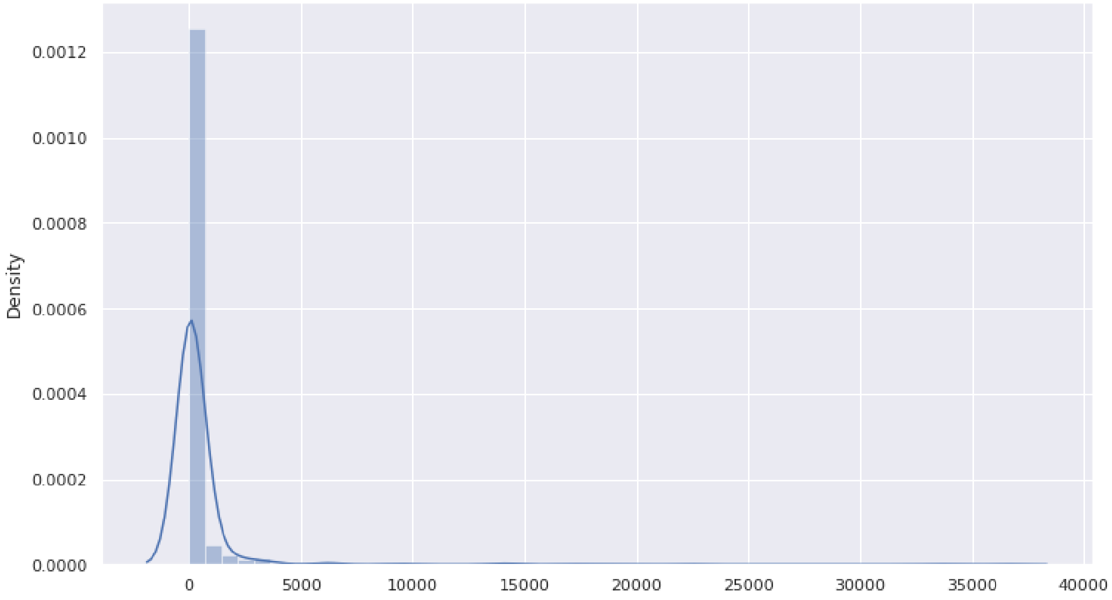


[65]:

```
top_10_complaints.plot(kind='hist')
plt.show()
```



[66]:
```
sns.distplot(complaint_by_city)
plt.show()
```



1.3 Create a DataFrame, df_new, which contains cities as columns and complaint types in rows

[67]:
```
df_new= service_request.pivot_table(index='Complaint Type', columns='City',␣
 ↪values='Unique Key', aggfunc='count')
df_new.head()
```

[67]:
```
City                        ARVERNE  ASTORIA  Astoria  BAYSIDE  BELLEROSE  \
Complaint Type
Agency Issues                   NaN      NaN      NaN      NaN        NaN
Animal Abuse                   46.0    170.0      NaN     53.0       15.0
Animal in a Park                NaN      NaN      NaN      NaN        NaN
Bike/Roller/Skate Chronic       NaN     16.0      NaN      NaN        1.0
Blocked Driveway               50.0   3436.0    159.0    514.0      138.0

City                        BREEZY POINT    BRONX  BROOKLYN  CAMBRIA HEIGHTS  \
Complaint Type
Agency Issues                        NaN      NaN       NaN              NaN
Animal Abuse                         2.0   1971.0    3191.0             15.0
Animal in a Park                     NaN      NaN       NaN              NaN
Bike/Roller/Skate Chronic            NaN     22.0     124.0              NaN
Blocked Driveway                     3.0  17061.0   36444.0            177.0
```

```
City                       CENTRAL PARK  …  SOUTH OZONE PARK  \
Complaint Type                           …
Agency Issues                       NaN  …               NaN
Animal Abuse                        NaN  …              74.0
Animal in a Park                    NaN  …               NaN
Bike/Roller/Skate Chronic           NaN  …               1.0
Blocked Driveway                    NaN  …            1202.0

City                       SOUTH RICHMOND HILL  SPRINGFIELD GARDENS  \
Complaint Type
Agency Issues                              NaN                  NaN
Animal Abuse                              40.0                 42.0
Animal in a Park                           NaN                  NaN
Bike/Roller/Skate Chronic                  1.0                  NaN
Blocked Driveway                        1946.0                330.0

City                       STATEN ISLAND  SUNNYSIDE  Unknown City  WHITESTONE  \
Complaint Type
Agency Issues                        NaN        NaN           8.0         NaN
Animal Abuse                       786.0       40.0           1.0        43.0
Animal in a Park                     NaN        NaN           NaN         NaN
Bike/Roller/Skate Chronic           10.0        2.0           2.0         4.0
Blocked Driveway                  2845.0      278.0          86.0       279.0

City                       WOODHAVEN  WOODSIDE  Woodside
Complaint Type
Agency Issues                    NaN       NaN       NaN
Animal Abuse                    57.0     111.0       NaN
Animal in a Park                 NaN       NaN       NaN
Bike/Roller/Skate Chronic        2.0       5.0       NaN
Blocked Driveway              1363.0    2037.0      27.0

[5 rows x 54 columns]
```

[68]:

```
df_new.shape
```

[68]:

```
(23, 54)
```

[69]:

```
df_new.describe()
```

[69]:

```
City     ARVERNE      ASTORIA     Astoria     BAYSIDE  BELLEROSE  \
count  17.000000    20.000000    5.000000   15.000000  18.000000
mean   15.235294   399.550000  181.000000  103.333333  27.055556
std    20.477211   845.711751  117.819778  201.856620  48.352228
min     1.000000     2.000000   14.000000    1.000000   1.000000
25%     1.000000     8.750000  145.000000    2.500000   1.000000
```

```
50%     2.000000    50.000000  159.000000    9.000000    1.500000
75%    29.000000   279.250000  277.000000   50.000000   14.500000
max    62.000000  3436.000000  310.000000  638.000000  138.000000

City   BREEZY POINT         BRONX      BROOKLYN  CAMBRIA HEIGHTS  \
count      8.000000     20.000000     20.000000        11.000000
mean       3.875000   2458.150000   5942.050000        56.090909
std        5.026714   4492.085933  10819.292989        65.506419
min        1.000000     15.000000     49.000000         1.000000
25%        1.000000     46.500000    112.750000         6.500000
50%        2.500000    351.000000    760.500000        19.000000
75%        3.250000   2537.500000   6038.000000       106.500000
max       16.000000  17061.000000  36444.000000       177.000000

City   CENTRAL PARK  …  SOUTH OZONE PARK  SOUTH RICHMOND HILL  \
count      2.000000  …         19.000000            17.000000
mean      55.000000  …        140.421053           201.823529
std       70.710678  …        302.587052           477.398580
min        5.000000  …          1.000000             1.000000
25%       30.000000  …          2.000000             2.000000
50%       55.000000  …          5.000000            24.000000
75%       80.000000  …         89.500000            93.000000
max      105.000000  …       1202.000000          1946.000000

City   SPRINGFIELD GARDENS  STATEN ISLAND   SUNNYSIDE  Unknown City  \
count            17.000000      20.000000   17.000000     15.000000
mean             64.294118     766.600000   55.529412     44.933333
std             112.246807    1493.705756   86.579528     82.236998
min               1.000000       6.000000    1.000000      1.000000
25%               2.000000      18.750000    3.000000      1.500000
50%               7.000000     132.500000   16.000000      7.000000
75%              42.000000     783.750000   53.000000     71.000000
max             330.000000    6223.000000  278.000000    312.000000

City   WHITESTONE    WOODHAVEN     WOODSIDE   Woodside
count   15.000000    16.000000    18.000000    5.00000
mean    91.266667   193.812500   242.000000   33.20000
std    176.164885   387.778497   516.029525   51.68849
min      1.000000     1.000000     1.000000    2.00000
25%      2.000000     3.000000     5.750000    5.00000
50%     21.000000     8.500000    38.000000    8.00000
75%     39.000000   119.000000   226.000000   27.00000
max    631.000000  1363.000000  2037.000000  124.00000

[8 rows x 54 columns]
```

1.4 Visualize the Major types of Complaints in Each City

```
[70]: service_request['City'].unique()
```

```
[70]: array(['NEW YORK', 'ASTORIA', 'BRONX', 'ELMHURST', 'BROOKLYN',
             'KEW GARDENS', 'JACKSON HEIGHTS', 'MIDDLE VILLAGE', 'REGO PARK',
             'SAINT ALBANS', 'JAMAICA', 'SOUTH RICHMOND HILL', 'Unknown City',
             'RIDGEWOOD', 'HOWARD BEACH', 'FOREST HILLS', 'STATEN ISLAND',
             'OZONE PARK', 'RICHMOND HILL', 'WOODHAVEN', 'FLUSHING', 'CORONA',
             'QUEENS VILLAGE', 'OAKLAND GARDENS', 'HOLLIS', 'MASPETH',
             'EAST ELMHURST', 'SOUTH OZONE PARK', 'WOODSIDE', 'FRESH MEADOWS',
             'LONG ISLAND CITY', 'ROCKAWAY PARK', 'SPRINGFIELD GARDENS',
             'COLLEGE POINT', 'BAYSIDE', 'GLEN OAKS', 'FAR ROCKAWAY',
             'BELLEROSE', 'LITTLE NECK', 'CAMBRIA HEIGHTS', 'ROSEDALE',
             'SUNNYSIDE', 'WHITESTONE', 'ARVERNE', 'FLORAL PARK',
             'NEW HYDE PARK', 'CENTRAL PARK', 'BREEZY POINT', 'QUEENS',
             'Astoria', 'Long Island City', 'Woodside', 'East Elmhurst',
             'Howard Beach'], dtype=object)
```

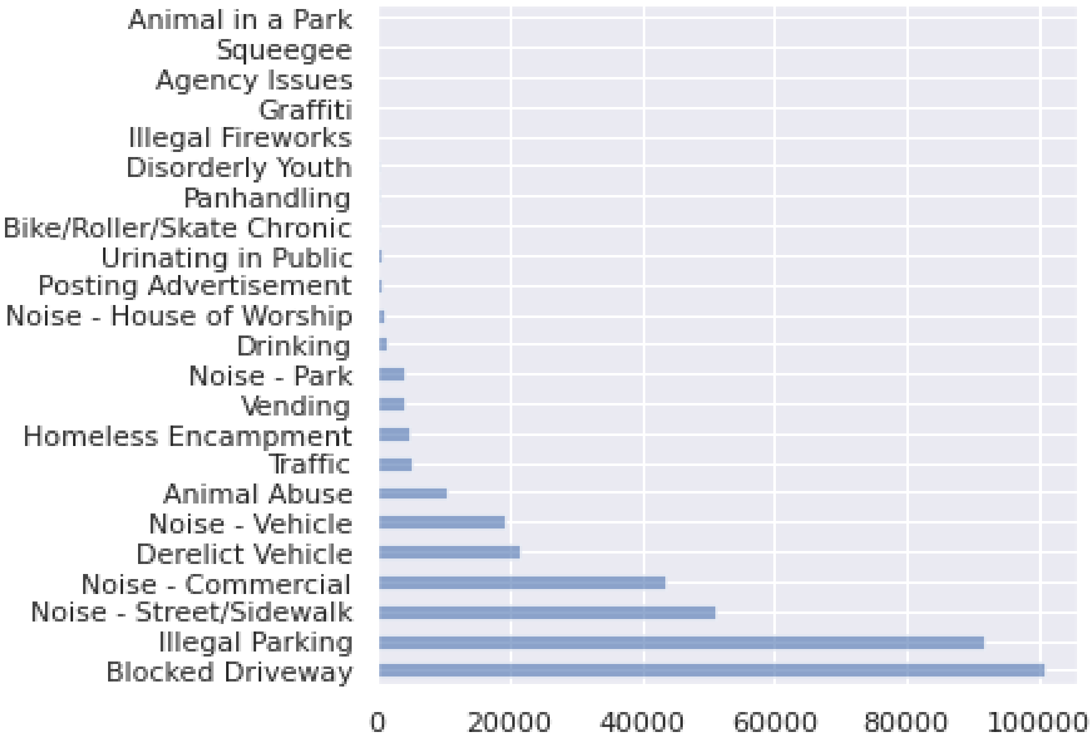
```
[71]: df_new.head()
```

```
[71]: City                       ARVERNE  ASTORIA  Astoria  BAYSIDE  BELLEROSE  \
      Complaint Type
      Agency Issues                  NaN      NaN      NaN      NaN        NaN
      Animal Abuse                  46.0    170.0      NaN     53.0       15.0
      Animal in a Park               NaN      NaN      NaN      NaN        NaN
      Bike/Roller/Skate Chronic      NaN     16.0      NaN      NaN        1.0
      Blocked Driveway              50.0   3436.0    159.0    514.0      138.0

      City                       BREEZY POINT    BRONX  BROOKLYN  CAMBRIA HEIGHTS  \
      Complaint Type
      Agency Issues                       NaN      NaN       NaN              NaN
      Animal Abuse                        2.0   1971.0    3191.0             15.0
      Animal in a Park                    NaN      NaN       NaN              NaN
      Bike/Roller/Skate Chronic           NaN     22.0     124.0              NaN
      Blocked Driveway                    3.0  17061.0   36444.0            177.0

      City                       CENTRAL PARK  …  SOUTH OZONE PARK  \
      Complaint Type                           …
      Agency Issues                       NaN  …               NaN
      Animal Abuse                        NaN  …              74.0
      Animal in a Park                    NaN  …               NaN
      Bike/Roller/Skate Chronic           NaN  …               1.0
      Blocked Driveway                    NaN  …            1202.0

      City                       SOUTH RICHMOND HILL  SPRINGFIELD GARDENS  \
      Complaint Type
```

```
Agency Issues                              NaN                  NaN
Animal Abuse                              40.0                 42.0
Animal in a Park                           NaN                  NaN
Bike/Roller/Skate Chronic                  1.0                  NaN
Blocked Driveway                        1946.0                330.0

City                       STATEN ISLAND  SUNNYSIDE  Unknown City  WHITESTONE  \
Complaint Type
Agency Issues                        NaN        NaN           8.0         NaN
Animal Abuse                       786.0       40.0           1.0        43.0
Animal in a Park                     NaN        NaN           NaN         NaN
Bike/Roller/Skate Chronic           10.0        2.0           2.0         4.0
Blocked Driveway                  2845.0      278.0          86.0       279.0

City                       WOODHAVEN  WOODSIDE  Woodside
Complaint Type
Agency Issues                    NaN       NaN       NaN
Animal Abuse                    57.0     111.0       NaN
Animal in a Park                 NaN       NaN       NaN
Bike/Roller/Skate Chronic        2.0       5.0       NaN
Blocked Driveway              1363.0    2037.0      27.0

[5 rows x 54 columns]
```

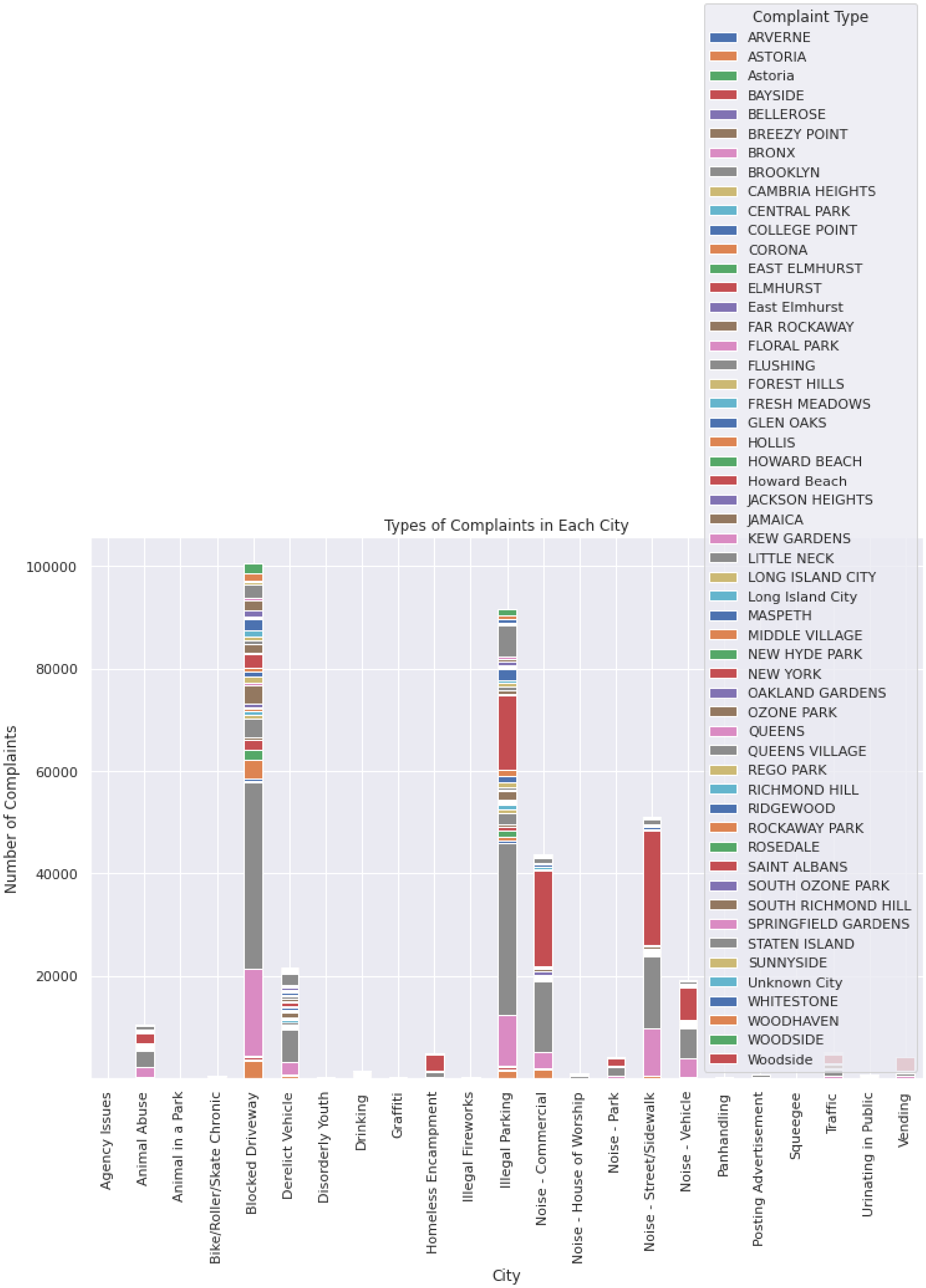
[89]:

```
service_request['Complaint Type'].value_counts().plot(kind='barh',alpha=0.
↪6,figsize=(5,5))
plt.show()
```



1.5 Draw another chart that shows the types of complaints in each city in a single chart, where different colors show the different types of complaints

[90]:
```
df_new.plot(kind='bar', stacked=True, figsize=(12, 8))
plt.xlabel('City')
plt.ylabel('Number of Complaints')
plt.title('Types of Complaints in Each City')
plt.legend(title='Complaint Type')
plt.show()
```


Types of Complaints in Each City
Number of Complaints
100000
80000
60000
40000
20000
City
Agency Issues
Animal Abuse
Animal in a Park
Bike/Roller/Skate Chronic
Blocked Driveway
Derelict Vehicle
Disorderly Youth
Drinking
Graffiti
Homeless Encampment
Illegal Fireworks
Illegal Parking
Noise - Commercial
Noise - House of Worship
Noise - Park
Noise - Street/Sidewalk
Noise - Vehicle
Panhandling
Posting Advertisement
Squeegee
Traffic
Urinating in Public
Vending
Complaint Type
ARVERNE
ASTORIA
Astoria
BAYSIDE
BELLEROSE
BREEZY POINT
BRONX
BROOKLYN
CAMBRIA HEIGHTS
CENTRAL PARK
COLLEGE POINT
CORONA
EAST ELMHURST
ELMHURST
East Elmhurst
FAR ROCKAWAY
FLORAL PARK
FLUSHING
FOREST HILLS
FRESH MEADOWS
GLEN OAKS
HOLLIS
HOWARD BEACH
Howard Beach
JACKSON HEIGHTS
JAMAICA
KEW GARDENS
LITTLE NECK
LONG ISLAND CITY
Long Island City
MASPETH
MIDDLE VILLAGE
NEW HYDE PARK
NEW YORK
OAKLAND GARDENS
OZONE PARK
QUEENS
QUEENS VILLAGE
REGO PARK
RICHMOND HILL
RIDGEWOOD
ROCKAWAY PARK
ROSEDALE
SAINT ALBANS
SOUTH OZONE PARK
SOUTH RICHMOND HILL
SPRINGFIELD GARDENS
STATEN ISLAND
SUNNYSIDE
Unknown City
WHITESTONE
WOODHAVEN
WOODSIDE
Woodside

1.6 Sort the complaint types based on the average Request_Closing_Timegrouping them for different locations

```
[91]: group_complaint =service_request.groupby('Complaint Type')
```

```
[92]: group_complaint['Complaint Type'].value_counts().nlargest(1)
```

```
[92]: Complaint Type    Complaint Type
      Blocked Driveway  Blocked Driveway    100618
      Name: Complaint Type, dtype: int64
```

```
[93]: # let us check whether the location type is null or non-null value
      service_request['Location Type'].isna().sum(axis=0)
```

```
[93]: 93
```

```
[94]: service_request.head()
```

```
[94]:    Unique Key         Created Date         Closed Date Agency  \
      0    32310363 2015-12-31 23:59:45 2016-01-01 00:55:15   NYPD
      1    32309934 2015-12-31 23:59:44 2016-01-01 01:26:57   NYPD
      2    32309159 2015-12-31 23:59:29 2016-01-01 04:51:03   NYPD
      3    32305098 2015-12-31 23:57:46 2016-01-01 07:43:13   NYPD
      4    32306529 2015-12-31 23:56:58 2016-01-01 03:24:42   NYPD

                             Agency Name           Complaint Type  \
      0  New York City Police Department  Noise - Street/Sidewalk
      1  New York City Police Department         Blocked Driveway
      2  New York City Police Department         Blocked Driveway
      3  New York City Police Department          Illegal Parking
      4  New York City Police Department          Illegal Parking

                           Descriptor    Location Type  Incident Zip  \
      0              Loud Music/Party  Street/Sidewalk       10034.0
      1                     No Access  Street/Sidewalk       11105.0
      2                     No Access  Street/Sidewalk       10458.0
      3  Commercial Overnight Parking  Street/Sidewalk       10461.0
      4              Blocked Sidewalk  Street/Sidewalk       11373.0

              Incident Address  … Road Ramp Bridge Highway Segment  \
      0    71 VERMILYEA AVENUE  …       NaN                     NaN
      1        27-07 23 AVENUE  …       NaN                     NaN
      2  2897 VALENTINE AVENUE  …       NaN                     NaN
      3    2940 BAISLEY AVENUE  …       NaN                     NaN
      4          87-14 57 ROAD  …       NaN                     NaN

       Garage Lot Name Ferry Direction Ferry Terminal Name   Latitude  Longitude  \
0             NaN              NaN                 NaN  40.865682 -73.923501
1             NaN              NaN                 NaN  40.775945 -73.915094
2             NaN              NaN                 NaN  40.870325 -73.888525
3             NaN              NaN                 NaN  40.835994 -73.828379
4             NaN              NaN                 NaN  40.733060 -73.874170

                                   Location    time_elapsed  \
0   (40.86568153633767, -73.92350095571744) 0 days 00:55:30
1  (40.775945312321085, -73.91509393898605) 0 days 01:27:13
2  (40.870324522111424, -73.88852464418646) 0 days 04:51:34
3   (40.83599404683083, -73.82837939584206) 0 days 07:45:27
4  (40.733059618956815, -73.87416975810375) 0 days 03:27:44

  time_elapsed_seconds
0               3330.0
1               5233.0
2              17494.0
3              27927.0
4              12464.0

[5 rows x 55 columns]
```

[95]:

```
# let is fill 1707 nan value with "unknown Location"
service_request['Location Type'].fillna('Unknown Location', inplace=True)
service_request['Location Type'].unique()
```

[95]:

```
array(['Street/Sidewalk', 'Club/Bar/Restaurant', 'Store/Commercial',
       'House of Worship', 'Residential Building/House',
       'Residential Building', 'Park/Playground', 'Vacant Lot',
       'House and Store', 'Highway', 'Commercial', 'Roadway Tunnel',
       'Subway Station', 'Parking Lot', 'Bridge', 'Unknown Location',
       'Park'], dtype=object)
```

[96]:

```
#Group dataset by major complaint type to display
grp_data = group_complaint.get_group('Blocked Driveway')
grp_data.shape
```

[96]:

```
(100618, 55)
```

[97]:

```
#fix those NAN with "Location Type" value instead
grp_data['Location Type'].fillna('Unknown Location', inplace =True)
```

[98]:

```
#Display the array for location
service_request['Location Type'].unique()
```

[98]:

```
array(['Street/Sidewalk', 'Club/Bar/Restaurant', 'Store/Commercial',
       'House of Worship', 'Residential Building/House',
        'Residential Building', 'Park/Playground', 'Vacant Lot',
        'House and Store', 'Highway', 'Commercial', 'Roadway Tunnel',
        'Subway Station', 'Parking Lot', 'Bridge', 'Unknown Location',
        'Park'], dtype=object)
```

[99]:
```
# display complaint type column and city type column together
service_request[['Complaint Type','City']]
```

[99]:
```
                 Complaint Type              City
0       Noise - Street/Sidewalk          NEW YORK
1              Blocked Driveway           ASTORIA
2              Blocked Driveway             BRONX
3               Illegal Parking             BRONX
4               Illegal Parking          ELMHURST
…                           …                …
364553          Illegal Parking         WOODHAVEN
364554          Noise - Vehicle             BRONX
364555  Noise - Street/Sidewalk          NEW YORK
364556         Blocked Driveway             BRONX
364557         Blocked Driveway  SOUTH OZONE PARK

[362138 rows x 2 columns]
```

Now lets find the mean value(average request closing time i.e average time elapsed between the created date and closed date), and group thnr for different location

[100]:
```
service_request.columns
```

[100]:
```
Index(['Unique Key', 'Created Date', 'Closed Date', 'Agency', 'Agency Name',
       'Complaint Type', 'Descriptor', 'Location Type', 'Incident Zip',
       'Incident Address', 'Street Name', 'Cross Street 1', 'Cross Street 2',
       'Intersection Street 1', 'Intersection Street 2', 'Address Type',
       'City', 'Landmark', 'Facility Type', 'Status', 'Due Date',
       'Resolution Description', 'Resolution Action Updated Date',
       'Community Board', 'Borough', 'X Coordinate (State Plane)',
       'Y Coordinate (State Plane)', 'Park Facility Name', 'Park Borough',
       'School Name', 'School Number', 'School Region', 'School Code',
       'School Phone Number', 'School Address', 'School City', 'School State',
       'School Zip', 'School Not Found', 'School or Citywide Complaint',
       'Vehicle Type', 'Taxi Company Borough', 'Taxi Pick Up Location',
       'Bridge Highway Name', 'Bridge Highway Direction', 'Road Ramp',
       'Bridge Highway Segment', 'Garage Lot Name', 'Ferry Direction',
       'Ferry Terminal Name', 'Latitude', 'Longitude', 'Location',
       'time_elapsed', 'time_elapsed_seconds'],
      dtype='object')
```

[101]:
```
pd.DataFrame(service_request.groupby("Location Type")["time_elapsed_seconds"].
 ↪mean()).sort_values("time_elapsed_seconds")
```

[101]:
```
                              time_elapsed_seconds
Location Type
Subway Station                        8.707200e+03
Club/Bar/Restaurant                   1.100953e+04
House of Worship                      1.140317e+04
Store/Commercial                      1.157573e+04
Highway                               1.226234e+04
Park/Playground                       1.239568e+04
Bridge                                1.376750e+04
Street/Sidewalk                       1.566325e+04
Residential Building                  1.603562e+04
Commercial                            1.623899e+04
Roadway Tunnel                        1.700916e+04
House and Store                       1.750501e+04
Parking Lot                           1.779160e+04
Unknown Location                      1.781462e+04
Residential Building/House            1.801399e+04
Vacant Lot                            2.427372e+04
Park                                  1.212634e+06
```

[102]:
```
pd.DataFrame(service_request.groupby("City")["time_elapsed_seconds"].mean()).
 ↪sort_values("time_elapsed_seconds")
```

[102]:
```
                     time_elapsed_seconds
City
ARVERNE                       8270.436293
ROCKAWAY PARK                 8376.174489
LITTLE NECK                   9301.886236
OAKLAND GARDENS               9374.410042
BAYSIDE                       9603.778710
FAR ROCKAWAY                  9671.584109
NEW YORK                     10520.947335
FLUSHING                     10646.788672
FOREST HILLS                 11045.858153
WHITESTONE                   11278.588020
CORONA                       11340.926979
COLLEGE POINT                11423.626943
JACKSON HEIGHTS              11453.122090
ELMHURST                     11646.878091
FRESH MEADOWS                12031.295390
REGO PARK                    12147.728279
BREEZY POINT                 12311.870968
EAST ELMHURST                12408.088839
CENTRAL PARK                 12415.281818
```

```
STATEN ISLAND              13683.176559
BROOKLYN                   14195.838759
Howard Beach               14505.000000
Astoria                    14547.138122
Long Island City           14723.335294
ASTORIA                    15914.190089
RIDGEWOOD                  16097.132843
SAINT ALBANS               16262.446036
East Elmhurst              16417.833333
Woodside                   16887.337349
KEW GARDENS                16999.186508
JAMAICA                    18328.541083
SOUTH OZONE PARK           18496.982759
SOUTH RICHMOND HILL        19081.228213
RICHMOND HILL              19304.943849
WOODHAVEN                  19306.596259
MIDDLE VILLAGE             19397.429507
OZONE PARK                 19698.548752
MASPETH                    19739.862368
HOLLIS                     19923.685621
HOWARD BEACH               20817.576923
BRONX                      21187.635986
LONG ISLAND CITY           22035.387640
SUNNYSIDE                  22844.657839
WOODSIDE                   23389.603535
NEW HYDE PARK              25403.790698
GLEN OAKS                  30099.207756
SPRINGFIELD GARDENS        30625.277219
CAMBRIA HEIGHTS            32572.987034
Unknown City               33399.753709
ROSEDALE                   34151.684693
BELLEROSE                  34570.416838
QUEENS VILLAGE             35635.228343
FLORAL PARK                36588.729592
QUEENS                     43030.270270
```

Above data shows that based on request time for closing files in different cities. let us find the percentage also for each city.

[103]:
```
print("Total Number of Concerns : ",len(service_request),"\n")
print("Percentage of Requests took less than 100 hour to get solved   :␣
 ↪",round((len(service_request)-(service_request["time_elapsed_seconds"]>100).
 ↪sum())/len(service_request)*100,2),"%")
print("Percentage of Requests took less than 1000 hour to get solved :␣
 ↪",round((len(service_request)-(service_request["time_elapsed_seconds"]>1000).
 ↪sum())/len(service_request)*100,2),"%")
```

```
Total Number of Concerns :  362138
```

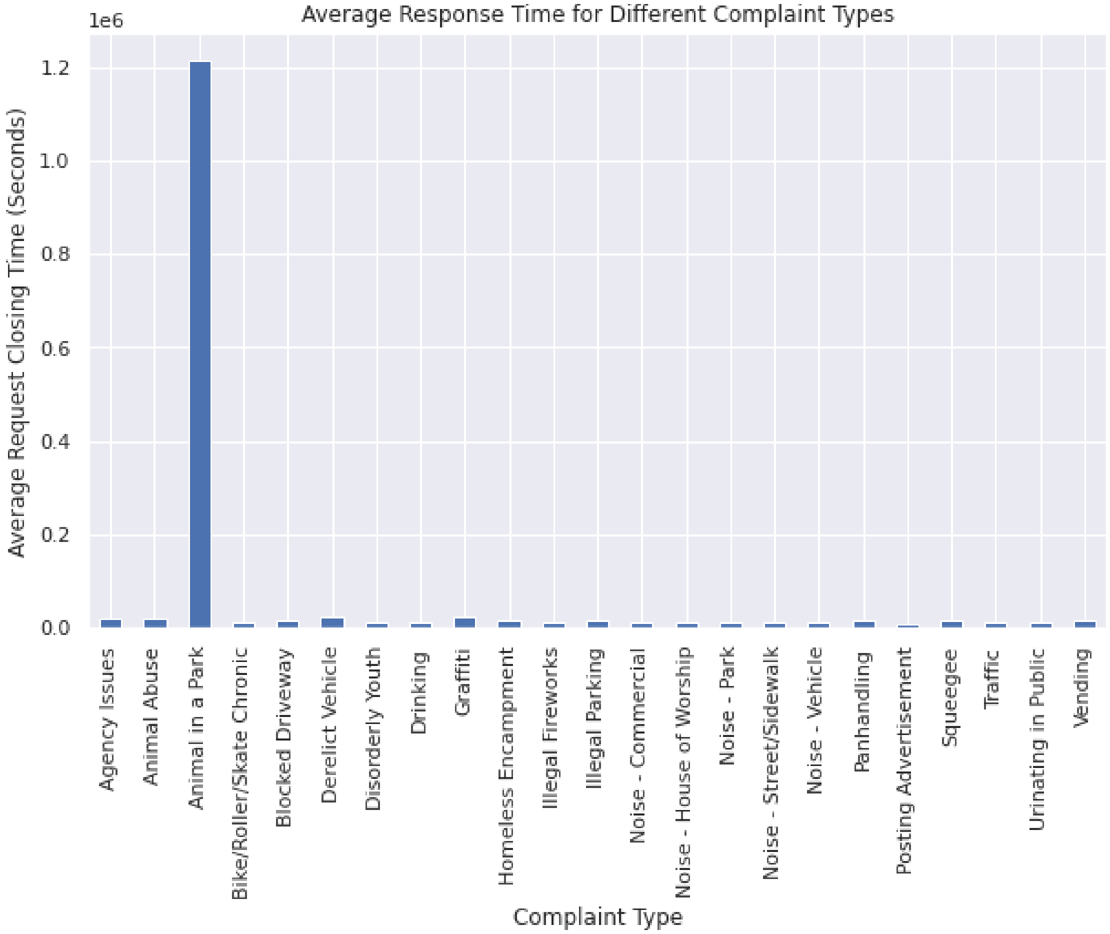
```
Percentage of Requests took less than 100 hour to get solved   :  0.0 %
Percentage of Requests took less than 1000 hour to get solved :  3.49 %
```

1.7 Visualize the average of Request_Closing_Time

we are making before the eliminating any larger null values

[104]:

```
service_request.groupby('Complaint Type')['time_elapsed_seconds'].mean().
 ↪plot(kind='bar', figsize=(10, 6))
plt.xlabel('Complaint Type')
plt.ylabel('Average Request Closing Time (Seconds)')
plt.title('Average Response Time for Different Complaint Types')
plt.show()
```



In the above graph we have obseverd that mostly the complaint type is Animal in the park

1.8 See whether the average response time across different complaint types is similar (overall)

let us handle the missing value firstly find the missing value and tret them accordingly. if the missig values are of larger percentange then will drop them for better visualizaton

[105]:

```
new_df= (service_request[service_request.columns].isna().sum(axis=0)/
 ↪service_request.shape[0])*100
print(new_df)
print("Old DataFrame for response time Shape :",service_request.shape)
print("New DataFrame for average response time Shape : ",new_df.shape)
```

```
Unique Key                         0.000000
Created Date                       0.000000
Closed Date                        0.000000
Agency                             0.000000
Agency Name                        0.000000
Complaint Type                     0.000000
Descriptor                         1.793239
Location Type                      0.000000
Incident Zip                       0.186393
Incident Address                  14.270527
Street Name                       14.270527
Cross Street 1                    15.277049
Cross Street 2                    15.313775
Intersection Street 1             86.021627
Intersection Street 2             86.055868
Address Type                       0.256532
City                               0.000000
Landmark                          99.896448
Facility Type                      0.004694
Status                             0.000000
Due Date                           0.000000
Resolution Description             0.000000
Resolution Action Updated Date     0.000000
Community Board                    0.000000
Borough                            0.000000
X Coordinate (State Plane)         0.471367
Y Coordinate (State Plane)         0.471367
Park Facility Name                 0.000000
Park Borough                       0.000000
School Name                        0.000000
School Number                      0.000000
School Region                      0.000276
School Code                        0.000276
School Phone Number                0.000000
School Address                     0.000000
School City                        0.000000
School State                        0.000000
School Zip                          0.000276
School Not Found                    0.000000
School or Citywide Complaint      100.000000
Vehicle Type                      100.000000
Taxi Company Borough              100.000000
Taxi Pick Up Location             100.000000
Bridge Highway Name                99.917987
Bridge Highway Direction           99.917987
Road Ramp                          99.927652
Bridge Highway Segment             99.927652
Garage Lot Name                   100.000000
Ferry Direction                   100.000000
Ferry Terminal Name               100.000000
Latitude                            0.471367
Longitude                           0.471367
Location                            0.471367
time_elapsed                        0.000000
time_elapsed_seconds                0.000000
dtype: float64
Old DataFrame for response time Shape : (362138, 55)
New DataFrame for average response time Shape :  (55,)
```

[106]:

```
new_df_1= service_request
new_df_1.head()
```

[106]:

```
   Unique Key        Created Date         Closed Date Agency  \
0    32310363 2015-12-31 23:59:45 2016-01-01 00:55:15   NYPD
1    32309934 2015-12-31 23:59:44 2016-01-01 01:26:57   NYPD
2    32309159 2015-12-31 23:59:29 2016-01-01 04:51:03   NYPD
3    32305098 2015-12-31 23:57:46 2016-01-01 07:43:13   NYPD
4    32306529 2015-12-31 23:56:58 2016-01-01 03:24:42   NYPD

                       Agency Name           Complaint Type  \
0  New York City Police Department  Noise - Street/Sidewalk
1  New York City Police Department         Blocked Driveway
2  New York City Police Department         Blocked Driveway
3  New York City Police Department          Illegal Parking
4  New York City Police Department          Illegal Parking

                     Descriptor    Location Type  Incident Zip  \
0              Loud Music/Party  Street/Sidewalk       10034.0
1                     No Access  Street/Sidewalk       11105.0
2                     No Access  Street/Sidewalk       10458.0
3  Commercial Overnight Parking  Street/Sidewalk       10461.0
4              Blocked Sidewalk  Street/Sidewalk       11373.0
```

```
        Incident Address  … Road Ramp Bridge Highway Segment  \
0    71 VERMILYEA AVENUE  …        NaN                    NaN
1        27-07 23 AVENUE  …        NaN                    NaN
2  2897 VALENTINE AVENUE  …        NaN                    NaN
3    2940 BAISLEY AVENUE  …        NaN                    NaN
4          87-14 57 ROAD  …        NaN                    NaN

  Garage Lot Name Ferry Direction Ferry Terminal Name   Latitude  Longitude  \
0             NaN             NaN                 NaN  40.865682 -73.923501
1             NaN             NaN                 NaN  40.775945 -73.915094
2             NaN             NaN                 NaN  40.870325 -73.888525
3             NaN             NaN                 NaN  40.835994 -73.828379
4             NaN             NaN                 NaN  40.733060 -73.874170

                                   Location     time_elapsed  \
0   (40.86568153633767, -73.92350095571744) 0 days 00:55:30
1  (40.775945312321085, -73.91509393898605) 0 days 01:27:13
2  (40.870324522111424, -73.88852464418646) 0 days 04:51:34
3   (40.83599404683083, -73.82837939584206) 0 days 07:45:27
4  (40.733059618956815, -73.87416975810375) 0 days 03:27:44

  time_elapsed_seconds
0               3330.0
1               5233.0
2              17494.0
3              27927.0
4              12464.0

[5 rows x 55 columns]
```

[107]:

```
rem=[]
for x in new_df_1.columns.tolist():
    if new_df_1[x].nunique()<=3:
        print(x+ " "*10+" : ",new_df_1[x].unique())
        rem.append(x)
```

```
Agency           :  ['NYPD']
Agency Name           :  ['New York City Police Department' 'Internal Affairs
Bureau']
Facility Type             :  ['Precinct' nan]
Status           :  ['Closed' 'Assigned']
Park Facility Name           :  ['Unspecified' 'Alley Pond Park - Nature
Center']
School Name           :  ['Unspecified' 'Alley Pond Park - Nature Center']
School Number             :  ['Unspecified' 'Q001']
School Region             :  ['Unspecified' nan]
School Code           :  ['Unspecified' nan]
School Phone Number           :  ['Unspecified' '7182176034']
School Address           :  ['Unspecified' 'Grand Central Parkway, near the
soccer field']
School City           :  ['Unspecified' 'QUEENS']
School State           :  ['Unspecified' 'NY']
School Zip           :  ['Unspecified' nan]
School Not Found           :  ['N']
School or Citywide Complaint           :  [nan]
Vehicle Type           :  [nan]
Taxi Company Borough           :  [nan]
Taxi Pick Up Location           :  [nan]
Road Ramp           :  [nan 'Roadway' 'Ramp']
Garage Lot Name           :  [nan]
Ferry Direction           :  [nan]
Ferry Terminal Name           :  [nan]
```

[108]:
```
new_df_1.drop(rem,axis=1,inplace=True)
new_df_1.shape
```

[108]:
```
(362138, 32)
```

[109]:
```
new_df_1.columns
```

[109]:
```
Index(['Unique Key', 'Created Date', 'Closed Date', 'Complaint Type',
       'Descriptor', 'Location Type', 'Incident Zip', 'Incident Address',
       'Street Name', 'Cross Street 1', 'Cross Street 2',
       'Intersection Street 1', 'Intersection Street 2', 'Address Type',
       'City', 'Landmark', 'Due Date', 'Resolution Description',
       'Resolution Action Updated Date', 'Community Board', 'Borough',
       'X Coordinate (State Plane)', 'Y Coordinate (State Plane)',
       'Park Borough', 'Bridge Highway Name', 'Bridge Highway Direction',
       'Bridge Highway Segment', 'Latitude', 'Longitude', 'Location',
       'time_elapsed', 'time_elapsed_seconds'],
      dtype='object')
```

[110]:
```
#Remove columns that are not needed for our analysis

rem1=["Unique Key","Incident Address","Descriptor","Street Name","Cross Street␣
 ↪1","Cross Street 2","Due Date","Resolution Description","Resolution Action␣
 ↪Updated Date","Community Board","X Coordinate (State Plane)","Y Coordinate␣
 ↪(State Plane)","Park Borough","Latitude","Longitude","Location"]

new_df_1.drop(rem1,axis=1,inplace=True)
```

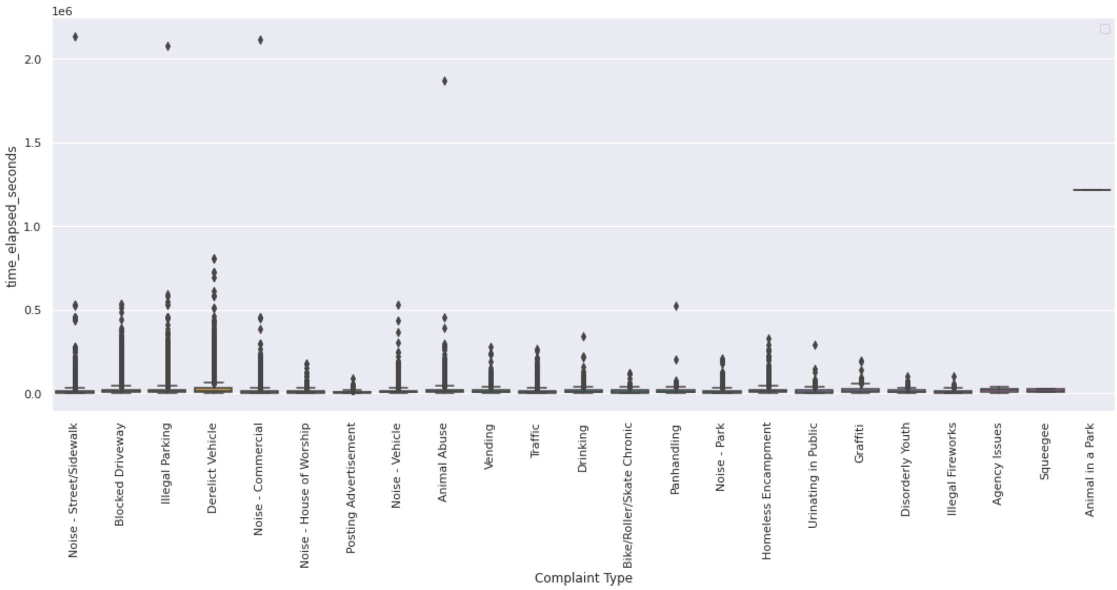
[111]:
```
new_df_1.columns
```

```
[111]: Index(['Created Date', 'Closed Date', 'Complaint Type', 'Location Type',
              'Incident Zip', 'Intersection Street 1', 'Intersection Street 2',
              'Address Type', 'City', 'Landmark', 'Borough', 'Bridge Highway Name',
              'Bridge Highway Direction', 'Bridge Highway Segment', 'time_elapsed',
              'time_elapsed_seconds'],
             dtype='object')
```

Visualize the average of Request_Closing_Time in the new data where we have already removed the data whose null % was higher than 50

```
[113]: g=sns.catplot(x="Complaint␣
        ↪Type",y="time_elapsed_seconds",kind="box",data=new_df_1,)
       g.fig.set_figheight(8)
       g.fig.set_figwidth(15)
       plt.xticks(rotation=90)
       plt.tight_layout()
       plt.legend()

       plt.show()
```

```
No artists with labels found to put in legend.  Note that artists whose label
start with an underscore are ignored when legend() is called with no argument.
```



1.9 Identify the significant variables by performing statistical analysis using p-values

[114]:

```
import scipy
from scipy import stats
```

To find the p value let us take either t test or ANOVA test, for multivarible test lets us follow Anova for better result

[115]:

```
anova_df=pd.DataFrame()
anova_df["Request_Closing_Time"]=new_df_1["time_elapsed_seconds"]
anova_df["Complaint"]=new_df_1["Complaint Type"]

anova_df.dropna(inplace=True)
anova_df.head()
```

[115]:

```
   Request_Closing_Time                 Complaint
0                3330.0  Noise - Street/Sidewalk
1                5233.0         Blocked Driveway
2               17494.0         Blocked Driveway
3               27927.0          Illegal Parking
4               12464.0          Illegal Parking
```

Since p value for the Complaint is less that 0.01 thus we accept alternate hypothesis i.e there is significant difference in the mean response time w.r.t different type of complaint.

*H0:Complaint Type and Location Type are independent

*H1:Complaint Type and Location Type are related

[116]:

```
from sklearn.impute import SimpleImputer
from sklearn import preprocessing
from sklearn.preprocessing import LabelEncoder
import statsmodels.api as sm
from statsmodels.formula.api import ols
from scipy.stats import chi2_contingency
```

Let us use the feature selection for encoding the categorical data in anova_df

[118]:

```
lm=ols("Request_Closing_Time~Complaint",data=anova_df).fit()
table=sm.stats.anova_lm(lm)
table
```

[118]:

```
                 df        sum_sq       mean_sq           F  PR(>F)
Complaint      22.0  5.354308e+12  2.433777e+11  565.212053     0.0
Residual   362115.0  1.559250e+14  4.305953e+08         NaN     NaN
```

2 let us find the chi__ square for the complaint and location for understanding the pvalue

[119]:
```
chi_sq=pd.DataFrame()
chi_sq["Location Type"]=new_df_1["Location Type"]
chi_sq["Complaint Type"]=new_df_1["Complaint Type"]

chi_sq.dropna(inplace=True)
chi_sq
```

[119]:
```
          Location Type            Complaint Type
0       Street/Sidewalk  Noise - Street/Sidewalk
1       Street/Sidewalk         Blocked Driveway
2       Street/Sidewalk         Blocked Driveway
3       Street/Sidewalk          Illegal Parking
4       Street/Sidewalk          Illegal Parking
…                 …                        …
364553  Street/Sidewalk          Illegal Parking
364554  Street/Sidewalk          Noise - Vehicle
364555  Street/Sidewalk  Noise - Street/Sidewalk
364556  Street/Sidewalk         Blocked Driveway
364557  Street/Sidewalk         Blocked Driveway

[362138 rows x 2 columns]
```

[120]:
```
chi2_crosstab =pd.crosstab(new_df_1['Location Type'], new_df_1['Complaint␣
 ↪Type'])
chi2_crosstab
```

[120]:
```
Complaint Type            Agency Issues  Animal Abuse  Animal in a Park  \
Location Type
Bridge                                0             0                 0
Club/Bar/Restaurant                   0             0                 0
Commercial                            0           108                 0
Highway                               0             0                 0
House and Store                       0           245                 0
House of Worship                      0             0                 0
Park                                  0             0                 1
Park/Playground                       0           143                 0
Parking Lot                           0           160                 0
Residential Building                  0           533                 0
Residential Building/House            0          6722                 0
Roadway Tunnel                        0             0                 0
Store/Commercial                      0           716                 0
Street/Sidewalk                       0          1874                 0
Subway Station                        0            26                 0
```

```
Unknown Location                        8             3                 0
Vacant Lot                              0             0                 0

Complaint Type              Bike/Roller/Skate Chronic  Blocked Driveway  \
Location Type
Bridge                                              0                 0
Club/Bar/Restaurant                                 0                 0
Commercial                                          0                 0
Highway                                             0                 0
House and Store                                     0                 0
House of Worship                                    0                 0
Park                                                0                 0
Park/Playground                                     0                 0
Parking Lot                                         0                 0
Residential Building                                0                 0
Residential Building/House                         31                 0
Roadway Tunnel                                      0                 0
Store/Commercial                                   60                 0
Street/Sidewalk                                   384            100586
Subway Station                                      0                 0
Unknown Location                                    0                32
Vacant Lot                                          0                 0

Complaint Type              Derelict Vehicle  Disorderly Youth  Drinking  \
Location Type
Bridge                                     0                 0         0
Club/Bar/Restaurant                        0                 0       457
Commercial                                 0                 0         0
Highway                                   18                 0         0
House and Store                            0                 0         0
House of Worship                           0                 0         0
Park                                       0                 0         0
Park/Playground                            0                 0        99
Parking Lot                                0                 0         0
Residential Building                       0                 0         0
Residential Building/House                 0                93       296
Roadway Tunnel                             8                 0         0
Store/Commercial                           0                10        96
Street/Sidewalk                        21389               212       455
Subway Station                             0                 0         0
Unknown Location                           6                 0         1
Vacant Lot                                95                 0         0

Complaint Type              Graffiti  Homeless Encampment  …  \
Location Type                                              …
Bridge                             0                    2  …
Club/Bar/Restaurant                0                    0  …
```

```
Commercial                         0                    0  …
Highway                            0                   19  …
House and Store                    0                    0  …
House of Worship                   0                    0  …
Park                               0                    0  …
Park/Playground                    0                  357  …
Parking Lot                        0                    0  …
Residential Building               0                    0  …
Residential Building/House        73                 1149  …
Roadway Tunnel                     0                    4  …
Store/Commercial                  51                  585  …
Street/Sidewalk                   33                 2754  …
Subway Station                     0                    0  …
Unknown Location                   0                    7  …
Vacant Lot                         0                    0  …

Complaint Type              Noise - House of Worship  Noise - Park  \
Location Type
Bridge                                             0             0
Club/Bar/Restaurant                                0             0
Commercial                                         0             0
Highway                                            0             0
House and Store                                    0             0
House of Worship                                1066             0
Park                                               0             0
Park/Playground                                    0          4088
Parking Lot                                        0             0
Residential Building                               0             0
Residential Building/House                         0             0
Roadway Tunnel                                     0             0
Store/Commercial                                   0             0
Street/Sidewalk                                    0             0
Subway Station                                     0             0
Unknown Location                                   2             0
Vacant Lot                                         0             0

Complaint Type              Noise - Street/Sidewalk  Noise - Vehicle  \
Location Type
Bridge                                            0                0
Club/Bar/Restaurant                               0                0
Commercial                                        0                0
Highway                                           0                0
House and Store                                   0                0
House of Worship                                  0                0
Park                                              0                0
Park/Playground                                   0                0
Parking Lot                                       0                0
```

```
Residential Building                              0                0
Residential Building/House                        0                0
Roadway Tunnel                                    0                0
Store/Commercial                                  0                0
Street/Sidewalk                               51130            19298
Subway Station                                    0                0
Unknown Location                                  1                2
Vacant Lot                                        0                0

Complaint Type              Panhandling  Posting Advertisement  Squeegee  \
Location Type
Bridge                                0                      0         0
Club/Bar/Restaurant                   0                      0         0
Commercial                            0                      0         0
Highway                               0                      0         0
House and Store                       0                      0         0
House of Worship                      0                      0         0
Park                                  0                      0         0
Park/Playground                       6                      0         0
Parking Lot                           0                      7         0
Residential Building                  0                      0         0
Residential Building/House           20                     69         0
Roadway Tunnel                        0                      0         0
Store/Commercial                     66                     10         0
Street/Sidewalk                     233                    592         4
Subway Station                        0                      0         0
Unknown Location                      0                      0         0
Vacant Lot                            0                      0         0

Complaint Type              Traffic  Urinating in Public  Vending
Location Type
Bridge                            0                    0        0
Club/Bar/Restaurant               0                   25        0
Commercial                        0                    0        0
Highway                         227                    0        0
House and Store                   0                    0        0
House of Worship                  0                    0        0
Park                              0                    0        0
Park/Playground                   0                   41      106
Parking Lot                       0                    0        0
Residential Building              0                    0        0
Residential Building/House        0                  154      219
Roadway Tunnel                   31                    0        0
Store/Commercial                  0                   69      492
Street/Sidewalk                4933                  337     3366
Subway Station                    0                   14        0
Unknown Location                  2                    1        0
```

```
Vacant Lot                        0                    0        0

[17 rows x 23 columns]
```

Here lets check for null hypothesis and alternative hypothesis and varifies the dependencies of the location/area regarding the complaints arise.

[121]:
```python
stat, p, dof, expected = chi2_contingency(chi2_crosstab)

alpha = 0.05
if p <= alpha:
    print('Dependent (reject H0)')
else:
    print('Independent (H0 holds true)')
```

```
Dependent (reject H0)
```

2.1 Perform a Kruskal-WallisH test

1.Fail to reject H0: All sample distributions are equal

2. Reject H0: One or more sample distributions are not equal

[122]:
```python
from scipy.stats import kruskal

# Suppose we want to compare the Request_Closing_Time across different␣
 ↪complaint types
complaint_types = service_request['Complaint Type'].unique()
data_groups = [service_request[service_request['Complaint Type'] ==␣
 ↪complaint_type]['time_elapsed_seconds'] for complaint_type in␣
 ↪complaint_types]
h_statistic, p_value = kruskal(*data_groups)

# 7.1 Fail to reject H0: All sample distributions are equal
if p_value > 0.05:
    print("Fail to reject null hypothesis: All sample distributions are equal.")
# 7.2 Reject H0: One or more sample distributions are not equal
else:
    print("Reject null hypothesis: One or more sample distributions are not␣
 ↪equal.")
```

```
Reject null hypothesis: One or more sample distributions are not equal.
```

2.2 Present your Observation

Since p value for the chi square test is less than 0.05(LOS) we can conclude that Complaint Type is dependent on Location Type i.e specific type of complaint is raised from specific places.

Maximum Complaints are raised in road and parking (vehicle) related sectors On an average complains are closed in an span of 150 to 300 hours Transport and Road related issues are taking more time to get resolved as number of these cases are quite high. Number of cases from Borough goes as follows BROOKLYN > QUEENS > MANHATTAN > BRONX > STATEN ISLAND Complaint Type are Depentent on Location Type. Time taken for solving different complaint type are different.

[ ]:

```

```

[ ]:

```

```